

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 14, 2004, 16:55:45 ; Search time 557.336 Seconds  
(without alignments)  
3959.005 Million cell updates/sec

Title: US-09-807-802A-15  
Perfect score: 3251  
Sequence: 1 TAPGKRPEVQSPQEPDSSS.....NNGLYTEPRPIGTRYLTRPL 599

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2449703 seqs, 1841816367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09807802/runat\_11022004\_175610\_15987/app.query.fasta\_1.2389  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosu62  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09807802@cgn\_1\_1\_1100@runat\_11022004\_175610\_15987  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US05\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq2:  
14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:  
18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result Query  
No. Score Match Length DB ID Description  
-----

|    |        |       |      |    |                   |                   |
|----|--------|-------|------|----|-------------------|-------------------|
| 1  | 3251   | 100.0 | 4718 | 13 | US-10-291-583-6   | Sequence 6, Appli |
| 2  | 3251   | 100.0 | 7447 | 13 | US-10-216-870-11  | Sequence 11, Appl |
| 3  | 2820.5 | 86.8  | 3075 | 13 | US-10-291-583-26  | Sequence 26, Appl |
| 4  | 2815   | 86.6  | 4726 | 13 | US-10-291-583-8   | Sequence 8, Appli |
| 5  | 2808.5 | 86.4  | 3142 | 13 | US-10-291-583-25  | Sequence 25, Appl |
| 6  | 2801   | 86.2  | 8179 | 15 | US-10-205-942-5   | Sequence 5, Appli |
| 7  | 2764.5 | 85.0  | 4679 | 9  | US-09-804-898-1   | Sequence 1, Appli |
| 8  | 2764.5 | 85.0  | 4679 | 9  | US-09-945-681-10  | Sequence 10, Appl |
| 9  | 2764.5 | 85.0  | 4679 | 13 | US-10-136-819-6   | Sequence 6, Appli |
| 10 | 2764.5 | 85.0  | 4679 | 14 | US-10-038-972A-12 | Sequence 12, Appl |
| 11 | 2761.5 | 84.9  | 2205 | 13 | US-10-291-583-120 | Sequence 120, App |
| 12 | 2753.5 | 84.7  | 3122 | 13 | US-10-291-583-56  | Sequence 56, Appl |
| 13 | 2751.5 | 84.6  | 3123 | 13 | US-10-291-583-54  | Sequence 54, Appl |
| 14 | 2749.5 | 84.6  | 3113 | 13 | US-10-291-583-55  | Sequence 55, Appl |
| 15 | 2735.5 | 84.1  | 3123 | 13 | US-10-291-583-57  | Sequence 57, Appl |
| 16 | 2735   | 84.1  | 3122 | 13 | US-10-291-583-45  | Sequence 45, Appl |
| 17 | 2726   | 83.9  | 3117 | 13 | US-10-291-583-43  | Sequence 43, Appl |
| 18 | 2722.5 | 83.7  | 4680 | 13 | US-10-375-777-1   | Sequence 1, Appli |
| 19 | 2722.5 | 83.7  | 4680 | 14 | US-10-077-294-1   | Sequence 1, Appli |
| 20 | 2722.5 | 83.7  | 4680 | 14 | US-10-163-886-1   | Sequence 1, Appli |
| 21 | 2722.5 | 83.7  | 4680 | 15 | US-10-263-127-1   | Sequence 44, Appl |
| 22 | 2722   | 83.7  | 3121 | 13 | US-10-291-583-44  | Sequence 42, Appl |
| 23 | 2719   | 83.6  | 3122 | 13 | US-10-291-583-42  | Sequence 27, Appl |
| 24 | 2715   | 83.5  | 3128 | 13 | US-10-291-583-27  | Sequence 29, Appl |
| 25 | 2715   | 83.5  | 3197 | 13 | US-10-291-583-29  | Sequence 59, Appl |
| 26 | 2710   | 83.4  | 3129 | 13 | US-10-291-583-59  | Sequence 1, Appli |
| 27 | 2709.5 | 83.3  | 4721 | 13 | US-10-291-583-1   | Sequence 41, Appl |
| 28 | 2707   | 83.3  | 3123 | 13 | US-10-291-583-41  | Sequence 1, Appli |
| 29 | 2706.5 | 83.3  | 4675 | 10 | US-09-782-378A-1  | Sequence 2, Appli |
| 30 | 2706.5 | 83.3  | 4675 | 10 | US-09-782-378A-2  | Sequence 7, Appli |
| 31 | 2706.5 | 83.3  | 4675 | 13 | US-10-291-583-7   | Sequence 1, Appli |
| 32 | 2706.5 | 83.3  | 4675 | 15 | US-10-240-198-1   | Sequence 5, Appli |
| 33 | 2704.5 | 83.2  | 4385 | 13 | US-10-291-583-5   | Sequence 9, Appli |
| 34 | 2703   | 83.1  | 3098 | 13 | US-10-291-583-9   | Sequence 47, Appl |
| 35 | 2703   | 83.1  | 3128 | 13 | US-10-291-583-47  | Sequence 28, Appl |
| 36 | 2702   | 83.1  | 3128 | 13 | US-10-291-583-28  | Sequence 34, Appl |
| 37 | 2701.5 | 83.1  | 3106 | 13 | US-10-291-583-34  | Sequence 39, Appl |
| 38 | 2700   | 83.1  | 3084 | 13 | US-10-291-583-39  | Sequence 11, Appl |
| 39 | 2699   | 83.0  | 3121 | 13 | US-10-291-583-11  | Sequence 35, Appl |
| 40 | 2698   | 83.0  | 2489 | 13 | US-10-291-583-35  | Sequence 15, Appl |
| 41 | 2696.5 | 82.9  | 3127 | 13 | US-10-291-583-15  | Sequence 38, Appl |
| 42 | 2696.5 | 82.9  | 3276 | 13 | US-10-291-583-38  | Sequence 36, Appl |
| 43 | 2696   | 82.9  | 2495 | 13 | US-10-291-583-36  | Sequence 37, Appl |
| 44 | 2696   | 82.9  | 3098 | 13 | US-10-291-583-37  | Sequence 18, Appl |
| 45 | 2695   | 82.9  | 3106 | 13 | US-10-291-583-18  |                   |

ALIGNMENTS

RESULT 1

US-10-291-583-6  
; Sequence 6, Application US/10291583  
; Publication No. US20030138772A1  
; GENERAL INFORMATION:  
; APPLICANT: Gao, Guangping  
; APPLICANT: Wilson, James M.  
; APPLICANT: Alvira, Mauricio  
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus  
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1  
; FILE REFERENCE: UPN-02735USA  
; CURRENT APPLICATION NUMBER: US/10/291,583  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: US 60/350,607  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/341,117  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: US 60/377,066  
; PRIOR FILING DATE: 2002-05-01  
; PRIOR APPLICATION NUMBER: US 60/386,675  
; PRIOR FILING DATE: 2002-06-05  
; NUMBER OF SEQ ID NOS: 120  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6

```
; LENGTH: 4718
; TYPE: DNA
; ORGANISM: adeno-associated virus serotype 1
US-10-291-583-6

Alignment Scores:
Pred. NO.:      0      Length:      4718
Score:          3251.00  Matches:      599
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:      100.00%  Indels: 0
DB:              13      Gaps: 0

US-09-807-802A-15 (1-599) x US-10-291-583-6 (1-4718)

QY 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20
Db 2634 ACGGCTCCTGGAAGAAACGTCGCGTAGAGCAGTCGCCACAAGAGCCAGACTCCTCCTCG 2693
QY 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40
Db 2694 GGCATCGGCAAGACAGCCGAGCAGCCGCTGAAAGAGACTCAATTTTGGTCAGACTGGC 2753
QY 41 AspSerGluSerValProAspProGlnProLeuGlyGluProAlaThrProAlaAla 60
Db 2754 GACTCAGAGTCAGTCCCGATCCACAACCTCTCGGAGAACCTCCAGCAACCCCGCTGCT 2813
QY 61 ValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnGluGly 80
Db 2814 GTGGGACCTACTACATAGCTTCAGCGCGTGGCGCACCAATGSCAGACAATAACGAAGSC 2873
QY 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100
Db 2874 GCCGACGAGTGGGTAAATGCCCTCAGGAAATTTGGCATTTGCGATTCCACATGGCTGGGCGAC 2933
QY 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120
Db 2934 AGAGTCATCACCACCAGCACCCGACCTGGGCGCTTGCCCACTTACCAATAACCACTCTAC 2993
QY 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140
Db 2994 AAGCAATCTCCAGTGTCTTCAACGGGGCCAGCAACGACCACTACTTCGGCTACAGC 3053
QY 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160
Db 3054 ACCCCCTGGGGTATTTTGATTTCAACAGATTCCACTGCCACTTTTCCACCACGTCGACTGG 3113
QY 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180
Db 3114 CAGCGACTCATCAACAACAATTTGGGGATTCCGGCCCCAAGAGACTCAACTTCAACTCTTC 3173
QY 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200
Db 3174 AACATCCAAGTCAAGGAGGTCAACGACGAATGATGGCGTCACACCATCGCTAATAACCTT 3233
QY 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220
Db 3234 ACCAGCACGGTCAAGTCTTCTCGGACTTCGGAGTACCAGCTTCCGTAGCTCCTCGGCTCT 3293
QY 221 AlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGly 240
Db 3294 GCGCACCGGGTGCCTCCCTCCGTTCCCGCGGACGTTGTTTCATGATTCGCAATACGGC 3353
QY 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260
Db 3354 TACCTGACGCTCAACAATGGCAGCCCAAGCGTGGGACGTTTCATCTTTTACTGCCTGGAA 3413
QY 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280
Db 3414 TATTTCCCTTCTCAGATGCTGAGAACGGGCAACAACCTTACCTTCAGCTACACCTTTGAG 3473
QY 281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300
Db 3000 |||||
```

```
Db 3474 GAAGTGCCCTTTCCACAGCAGCTACGCGCACACAGCCAGAGCCCTGGACCGGCTGATGAATCCT 3533
QY 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320
Db 3534 CTCATCGACCAATACCTGTATTACCTGAACAGAACTCAAATCAGTCCGGAAGTGCCCAA 3593
QY 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340
Db 3594 AACAAAGGACTTGTGTTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTCAAGCCCAAAAC 3653
QY 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360
Db 3654 TGGCTACCTGGACCCCTGTATTATCGGCAGCAGCGCTTTCTAAAAACAAACACAGACAAC 3713
QY 361 AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380
Db 3714 AACAGCAATTTTACCTGGACTGGTGTCTTCAAAATATAACCTCAATGGGCGTGAATCCATC 3773
QY 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePhePromet 400
Db 3774 ATCAACCCCTGGCACTGTATGGCTTCACACAAAGACGACGACGAAAGACAAGTTCTTCCCATG 3833
QY 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420
Db 3834 AGCGGTGTCAATGATTTTGGAAAGAGAGCGCGGAGCTTCAAACACTGCATTGGACAAT 3893
QY 421 ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440
Db 3894 GTCATGATTACAGACGAAGAGGAAATTTAAAGCCACTAACCTGTGGCCACCGAAAGATTT 3953
QY 441 GlyThrValAlaValAsnPheGlnSerSerSerSerThrAspProAlaThrGlyAspValHis 460
Db 3954 GGGACCGTGGCAGTCAATTTCCAGAGCAGCAGCAGCAGCCTGCGACCGGAGATGTGCAT 4013
QY 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 480
Db 4014 GCTATGGGAGCATTAACCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGACGGTCCC 4073
QY 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500
Db 4074 ATTTGGGCCAAAATTCCTCACACAGATGGACACTTTTCAACCGTCTCCTCTTATGGCGGC 4133
QY 501 PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520
Db 4134 TTTGGACTCAAGAACCCGCTCCTCAGATCCTCATCAAAAAACACGCTGTCTCTGCGAAT 4193
QY 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540
Db 4194 CCTCCGGGGAGTTTTCAGCTACAAAGTTTGTCTTCAATCATCCCCAATACTCCACAGGA 4253
QY 541 GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro 560
Db 4254 CAAGTGAGTGTGGAATTTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGATCCC 4313
QY 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580
Db 4314 GAAGTGCAGTACACATCCAATTTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAC 4373
QY 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599
Db 4374 AATGGACTTTTATCTAGCTGAGCTCGCCCATTTGGCACCCGTTACCTTACCCGTCCTG 4430
```

RESULT 2

```
US-10-216-870-11
; Sequence 11, Application US/10216870
; Publication No. US20030148506A1
; GENERAL INFORMATION:
; APPLICANT: KOTIN, ROBERT M
; APPLICANT: URABE, MASASHI
; APPLICANT: DING, CHUAN-TIAN
; TITLE OF INVENTION: PRODUCTION OF ADENO-ASSOCIATED VIRUS IN INSECT CELLS
; FILE REFERENCE: 402133
; CURRENT APPLICATION NUMBER: US/10/216,870
```



```

; CURRENT FILING DATE: 2002-08-13
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 7447
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-216-870-11

Alignment Scores:
Pred. No.: 0 Length: 7447
Score: 3251.00 Matches: 599
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-807-802A-15 (1-599) x US-10-216-870-11 (1-7447)

QY 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20
Db 5032 ACGGCTCCTGGAAGAAACGTCGGTAGAGCAGTCGCCACAGAGCCAGACTCTCTCTCG 5091

QY 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40
Db 5092 GGCATCGGCAAGACAGCGCCAGAGCCCGCTAAAGAGAGACTCAATTTGGTCAGACTGGC 5151

QY 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60
Db 5152 GACTCAGAGTCAGTCCCGCATCCACAACCTCTCGGAGAACCTCCAGCAACCCCGCTGCT 5211

QY 61 ValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGly 80
Db 5212 GTGGACCTACTACATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAAATAACGAAGGC 5271

QY 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100
Db 5272 GCGACGGAGTGGGTAATGCCCTCAGGAAATTTGGCATTTGGCATTTCCACATGGCTGGCGGAC 5331

QY 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120
Db 5332 AGAGTCATCACCACAGCACCCGACCTGGGCCCTTGCCACCTACAAATAACCACTCTAC 5391

QY 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140
Db 5392 AAGCAATCTCCAGTGTCTTCAACGGGGGCCAGCAACGACCAACCCTACTTCCGGCTACAGC 5451

QY 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160
Db 5452 ACCCCCTGGGGTATTTTGATTTCAACAGATTCCACTGCCACTTTTCCACCACGTCAGTGG 5511

QY 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180
Db 5512 CAGCGACTCATCAACAACAATTTGGGGATTCCGGCCCAAGAGACTCAAACTTCAAACTCTTC 5571

QY 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200
Db 5572 AACATCCAAGTCAAGGAGTCCAGACGAATGATGGCGCCCAAGAGACTCAAACTTCAAACTCTTC 5631

QY 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220
Db 5632 ACCAGCAGGTTCAAGTCTTCTCGGACTCGGAGTACCACTCCGTCAGTCCCTCGGCTCT 5691

QY 221 AlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGly 240
Db 5692 GCGCACCAGGGCTGCCTCCCTCCGTTCCCGGGGACGCGTTCATGATTCGCGCAATACGGC 5751

QY 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260
Db 5752 TACCTGACGCTCAACAATGGCAGCCAGCCGTTGGGACGTTTCATCTCTTTTACTGCTGGAA 5811
```

```

QY 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280
Db 5812 TATTTCCCTTCTCAGATGCTGAGAACGGGCAACAACCTTTACCTTCAGCTACACCTTTGAG 5871

QY 281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300
Db 5872 GAAGTGCCTTTCCACAGCAGCTACGGCGACAGCCAGAGCCTGGACCGGCTGATGAATCCT 5931

QY 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320
Db 5932 CTCATCGACCAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAA 5991

QY 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340
Db 5992 AACAAGGACTTGTGTTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAAAC 6051

QY 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360
Db 6052 TGGCTACCTGGACCTGTTATCGGCAGCAGCGCGTTTCTAAACAAACAGACACAAC 6111

QY 361 AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380
Db 6112 AACAGCAATTTTACCTGGACTGGTGCTTCAAAATATAACCTCAATGGCGTGAAATCCATC 6171

QY 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400
Db 6172 ATCAACCCCTGGCACTGCTATGGCTCACAAAGACGACGAGACAAAGTCTTCTTCCCATG 6231

QY 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420
Db 6232 AGCGGTGTGTCATGATTTTGGAAAAGAGAGCGCGGAGCTTCAAAACACTGCAATGGACAAT 6291

QY 421 ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440
Db 6292 GTCATGATTACAGACGAAGAGGAAATTAAGCCACTAACCTGTGGCCACCGAAAGATTT 6351

QY 441 GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis 460
Db 6352 GGGACCGTGGCAGTCAATTTCCAGAGCAGCAGCAGACACCTTCCGACCCGGAGATGTGCAT 6411

QY 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 480
Db 6412 GCTATGGGAGCATTACCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCAGGGTCCC 6471

QY 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500
Db 6472 ATTTGGGCCAAAATTCCTCACACAGATGGACACTTTCACCCGTCTCCTCTTATGGCGGC 6531

QY 501 PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520
Db 6532 TTTGGACTCAAGAACCCGCTCTCTCAGATCCTCATCAAAAACACGCTGTTCCTGCGAAT 6591

QY 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540
Db 6592 CCTCCGGCGGAGTTTTCAGCTACAAAGTTTGTCTTTCATTCATCCCAATATCTCCACAGGA 6651

QY 541 GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro 560
Db 6652 CAAGTGAGTGTGGAATTTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCC 6711

QY 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580
Db 6712 GAAGTGCAGTACACATCCAATTAATGCAAAATCTGCCAAACGTTGATTTTACTGTGGACAAC 6771

QY 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599
Db 6772 AATGGACTTTTACTGAGCCTCGCCCCCATTTGGCACCCCGTTACCTTACCCGTCCTCCCTG 6828
```

RESULT 3  
US-10-291-583-26  
; Sequence 26, Application US/10291583  
; Publication No. US20030138772A1  
; GENERAL INFORMATION:

APPLICANT: Gao, Guangping  
APPLICANT: Wilson, James M.  
APPLICANT: Alvira, Mauricio  
TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus  
TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1  
FILE REFERENCE: UPN-02735USA  
CURRENT APPLICATION NUMBER: US/10/291,583  
CURRENT FILING DATE: 2002-11-12  
PRIOR APPLICATION NUMBER: US 60/350,607  
PRIOR FILING DATE: 2001-11-13  
PRIOR APPLICATION NUMBER: US 60/341,117  
PRIOR FILING DATE: 2001-12-17  
PRIOR APPLICATION NUMBER: US 60/377,066  
PRIOR FILING DATE: 2002-05-01  
PRIOR APPLICATION NUMBER: US 60/386,675  
PRIOR FILING DATE: 2002-06-05  
NUMBER OF SEQ ID NOS: 120  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 26  
LENGTH: 3075  
TYPE: DNA  
ORGANISM: new AAV serotype, clone H2  
US-10-291-583-26

Alignment Scores:  
Pred. No.: 8.28e-298 Length: 3075-  
Score: 2820.50 Matches: 506  
Percent Similarity: 91.32% Conservative: 41  
Best Local Similarity: 84.47% Mismatches: 51  
Query Match: 86.76% Indels: 1  
DB: 13 Gaps: 1

US-09-807-802A-15 (1-599) x US-10-291-583-26 (1-3075)

|    |      |  |      |
|----|------|--|------|
| QY | 1    | ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSer    | 20   |
| Db | 1214 | ACGGCTCCGGGAAAAAGAGCGGTAGAGCACTCTCCTGTGGAGCCAGACTCCTCCTCG    | 1273 |
| QY | 21   | GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly | 40   |
| Db | 1274 | GGAACCGGAAAAAGCGGCCAGCGGCTGCAAGAAAAAGATTAAATTTTGGTCAGACTGA   | 1333 |
| QY | 41   | AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla | 60   |
| Db | 1334 | GACGACACTCGTACCTGACCCCCAGCCTCTCGGACAGCCACAGCAGCCCCCTCTG      | 1393 |
| QY | 61   | ValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnGluGly    | 80   |
| Db | 1394 | CTGGGATCTACTACAAATGGCTACAGGCACTGGCGCACCAATGGCAGACAATAACGAG   | 1453 |
| QY | 81   | AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp | 100  |
| Db | 1454 | GCCGATGGAGTGGGTAATTCCTCAGGAATTTGGCATTGGCATTTCCCAATGGCTGGCG   | 1513 |
| QY | 101  | ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr | 120  |
| Db | 1514 | AGAGTCATCACACACAGCAGCCGAACTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCT    | 1573 |
| QY | 121  | LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer | 140  |
| Db | 1574 | AAGCAATCTCCAGCCAATCA--GGAGCCAGCAACGACCACTTGGCTTGGCTTGGCT     | 1630 |
| QY | 141  | ThrProTyrGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp | 160  |
| Db | 1631 | ACCCCTGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCCACCAGTGACTGG   | 1690 |
| QY | 161  | GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe | 180  |
| Db | 1691 | CAAAGACTCATCAACAACAACTGGGGATTCCGGCCCAAAAGACTCAACTTCAAGCTCTT  | 1750 |
| QY | 181  | AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnLeu    | 200  |
| Db | 1751 | AATATTCAAGTCAAGAGGTCACGCAGAAATGACGGTACGACGAGATTGCCAATAACCTT  | 1810 |

|    |      |   |      |
|----|------|---|------|
| QY | 201  | ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer    | 220  |
| Db | 1811 | ACCAGCACGGTTGAGTGTCTTACTGACTCGGAGTACCAGTCCCGTACGTCCTGGGCTCG     | 1870 |
| QY | 221  | AlaHisGlnGlyCysLeuProPheProPheProAlaAspValPheMetIleProGlnTyrGly | 240  |
| Db | 1871 | GCGCATCAAGGATGCTCCCGCGTTTCCAGCGGACGTCTTCTATGGTCCACAGTATGGA      | 1930 |
| QY | 241  | TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu    | 260  |
| Db | 1931 | TACCTCACCTGAAACAACGGGAGTCAAGCGGTAGGACGCTCTTCTCTTACTGCTGGAG      | 1990 |
| QY | 261  | TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu    | 280  |
| Db | 1991 | TACTTTCTCTCAGATGCTGCGTACTGGAACAACACTTTCAGTTTCAGCTACACTTTTGA     | 2050 |
| QY | 281  | GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro    | 300  |
| Db | 2051 | GACGTGCTTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGATCGGCTGATGAATCCT      | 2110 |
| QY | 301  | LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln    | 320  |
| Db | 2111 | CTGATCGACCACTGCTGTTATCTGAACAAGACACAACAATAAGTGAACACTCTTCAG       | 2170 |
| QY | 321  | AsnLysAspLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn       | 340  |
| Db | 2171 | CAGTCTCGGCTACTGTTTAGCCAAGCTGGACCAACCAACATGTCTCTTCAAGCTAAAAAC    | 2230 |
| QY | 341  | TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn    | 360  |
| Db | 2231 | TGGCTGCTGGACCTTGCTACAGACAGCAGCGCTGTCTCAAAACAGGCAACGACAAAC       | 2290 |
| QY | 361  | AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle    | 380  |
| Db | 2291 | AACAGCAACTTTCCTGGACTGCAGCTACAAAGTATCATCTAAATGGCCGGACTCGTTG      | 2350 |
| QY | 381  | IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePhePromet    | 400  |
| Db | 2351 | GTTAATCCAGGACCACTATGGCCAGTCAACAAGGATGACGAAGAAAGTTTTCCTCCATG     | 2410 |
| QY | 401  | SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn    | 420  |
| Db | 2411 | CATGGAACCTTGATATTGGTAAACAAGGAACAATAATGCCACGCGGATTTGGAAAT        | 2470 |
| QY | 421  | ValMetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe    | 440  |
| Db | 2471 | GTCATGATTACAGATGAAGAAGAAATCAGGGCCACCAATCCCGTGGCTACGGAGCAGTAC    | 2530 |
| QY | 441  | GlyThrValAlaValAsnPheGlnSerSerSerSerThrAspProAlaThrGlyAspValHis | 460  |
| Db | 2531 | GGGACTGTGTCAATAATTTGCAAAACTCAAACTCAAACTCAAACTCAAACTCAAACT       | 2590 |
| QY | 461  | AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro    | 480  |
| Db | 2591 | CGCCAAGGAGCGTTACCTGGTATGGTGTGGCAGGATCGAGACGTGTACCTGCAGGACCC     | 2650 |
| QY | 481  | IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly    | 500  |
| Db | 2651 | ATTGGGCCAAGATTCTCTCACACCGATGGACACTTTCATCTCTCTCCACTGATGGAGGT     | 2710 |
| QY | 501  | PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn       | 520  |
| Db | 2711 | TTTGGACTCAAAACACCCGCTCTCAGATCATGATCATGATCAAAAACACTCCCGTTCCAGC   | 2770 |
| QY | 521  | ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly    | 540  |
| Db | 2771 | CCTCCCAAAAATTTCAGTTCTGCCAAGTTTGTCTTCTTTCATCATCACAGTATTTCCACG    | 2830 |
| QY | 541  | GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro    | 560  |
| Db | 2831 | CAGGTACGCTGGAGATCGAGTGGGAGCTGCAGAGGAGAGACAGCAACGCTGGAATCCC      | 2890 |

```
QY 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580
|||::|||
Db 2891 GAAATTCAGTACACTTCCAACACTACAACAAAGTCTGTAATGTGGACTTTACTGTGGACACT 2950
|||::|||
QY 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599
|||::|||
Db 2951 AATGGTGTGATTTCAGAGCCTCGCCCAATTGGCACAGATACCTGACTCGTAATCTG 3007
|||::|||

RESULT 4
US-10-291-583-8
; Sequence 8, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1el Sequences Identifi
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 4726
; TYPE: DNA
; ORGANISM: adeno-associated virus serotype 3
US-10-291-583-8

Alignment Scores:
Pred. No.: 6.47e-297 Length: 4726
Score: 2815.00 Matches: 511
Percent Similarity: 91.00% Conservative: 35
Best Local Similarity: 85.17% Mismatches: 52
Query Match: 86.59% Indels: 2
DB: 13 Gaps: 2

US-09-807-802a-15 (1-599) x US-10-291-583-8 (1-4726)

QY 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20
|||::|||
Db 2620 ACGGCTCCTGGAAGAAGGGGCTGTAGATCAGTCTCCTCAGGAACCGGACTCATCT 2679
|||::|||
QY 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40
|||::|||
Db 2680 GGTGTGGCAAATCGGGCAAACAGCCTGCCAGAAAAGACTAAATTTGGTTCAGACTGGA 2739
|||::|||
QY 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60
|||::|||
Db 2740 GACTCAGAGTCAGTCCCAGACCTCAACCTCTCGGAGAACCAACCAGCAGCCCCCAAGT 2799
|||::|||
QY 61 ValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGly 80
|||::|||
Db 2800 TTGGGATCTAATACAATGGCTTCAGGCGGTGGCGCAACCAATGGCAGACAATAACGAGGT 2859
|||::|||
QY 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100
|||::|||
Db 2860 GCCGATGGAGTGGGTAAATTCCTCAGGAAATTTGGCATTTGCCAATCCCAATGGCTGGCGCAC 2919
|||::|||
QY 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120
|||::|||
Db 2920 AGAGTCATCACCACCAGCACCAACCACTGGGCCCTGCCCACTTACCAACCAATCTCTAC 2979
|||::|||
QY 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140
|||::|||
```

```
Db 2980 AAGCAAATCTCCAGCCAATCA--GGAGCTTCAAACGACAACCACTACTTTGGGTACAGC 3036
|||::|||
QY 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160
|||::|||
Db 3037 ACCCCTTGGGGGTATTTTGACTTTAACAGATTCCACTGCCACTTCTCACACACGTGACTGG 3096
|||::|||
QY 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180
|||::|||
Db 3097 CAGCGACTCATTAACAACAACACTGGGATTCCGGCCCAAGAACTCAGCTTCAAGCTCTTTC 3156
|||::|||
QY 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200
|||::|||
Db 3157 AACATCCAAGTTAGAGGGGTACGCGAGAACGATGGCAGCAGACTATTGCCAATAACCTT 3216
|||::|||
QY 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220
|||::|||
Db 3217 ACCAGCACGGTTCAAGTGTTTACGGACTCGAGTATCAGCTCCCGTACGTGCTCGGGTCG 3276
|||::|||
QY 221 AlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGly 240
|||::|||
Db 3277 GCGCACCAAGCTGTCTCCCGCGTTTCCAGCGGACGCTTTCATGTGTCTCCTCAGTATGGA 3336
|||::|||
QY 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260
|||::|||
Db 3337 TACCTCACCTGAACAACGGAAGTCAAGCGGTGGGACGCTCATCTTTTACTGCCTGGAG 3396
|||::|||
QY 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280
|||::|||
Db 3397 TACTTCCCTTCGCAGATGCTAAGGACTGGAAAATAACTTCCAATTTCAGTATACCTTCGAG 3456
|||::|||
QY 281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300
|||::|||
Db 3457 GATGTACCTTTTCACAGCAGCTACGCTCACAGCCAGAGTTTGGATCGCTTGATGAATCCT 3516
|||::|||
QY 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGln--AsnGlnSerGlySerAla 319
|||::|||
Db 3517 CTTATTGATCAGTATCTGTACTACCTGAACAGAACGGAACAACCTCTTGGAAACAACC 3576
|||::|||
QY 320 GlnAsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLys 339
|||::|||
Db 3577 AACCAATCACGGCTGCTTTTATAGCCAGGCTGGGCCTCAGTCTATGTCTTTCAGAGCCAGA 3636
|||::|||
QY 340 AsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsn 359
|||::|||
Db 3637 AATTGGCTACCTGGGCCCTGTCTACCGGCAACAGAGACTTTTCAAAGACTGCTTAACGACAAC 3696
|||::|||
QY 360 AsnAsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSer 379
|||::|||
Db 3697 AACAAACAGTAACCTTCTTGGACAGCGGCCAGCAAAATATCATCTCAATGGCGCGGACTCG 3756
|||::|||
QY 380 IleIleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePhePro 399
|||::|||
Db 3757 CTGGTGAATCCAGGACCAGCTATGGCCAGTCACAAGGACGATGAAGAAAATAATTTTCCCT 3816
|||::|||
QY 400 MetSerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAsp 419
|||::|||
Db 3817 ATGCACGGCAATCTAATATTTGGCAAAGAAGGACAAACGGCAAGTAACGCAGAAATTAGAT 3876
|||::|||
QY 420 AsnValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArg 439
|||::|||
Db 3877 AATGTAATGATTACGGATGAAGAAGAGATTTCGTACCACCAATCCTGTGGCAACAGAGCAG 3936
|||::|||
QY 440 PheGlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspVal 459
|||::|||
Db 3937 TATGGAACGTGGCAAAATAACTTGCAGAGCTCAAAATACAGCTCCCACGACTGGAACTGTC 3996
|||::|||
QY 460 HisAlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGly 479
|||::|||
Db 3997 AATCATCAGGGGGCTTACCTGGCATGGTGTGCAAGATCGTGACGTGTACCTTCAAGGA 4056
|||::|||
QY 480 ProIleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGly 499
|||::|||
Db 4057 CCTATCTGGGCAAGAGATTCTCTCACACGGATGGACACTTTCATCTCTTCTCTCTGATGGGA 4116
|||::|||
```





Db 671 GTCATGATTACAGATGAAGAATAATCAGGGCCACCCAATCCCGTGGCTACGGAGCAGTAC 612  
Qy GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis 460  
Db 611 GGGACTGTGTCAAATAATTGCAAAACTCAAACACTGGTCCAACTACTGGAACGTCAAT 552  
Qy AlaMetGlyAlaLeuProGlyMetValTrrPglNAspArgAspValTyrLeuGlnGlyPro 480  
Db 551 CACCAAGGAGCGTTACCTGGTATGGTGTGGCAGGATCGAGACGTGTACCTGCAGGGACCC 492  
Qy IleTrrPAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500  
Db 491 ATTGGGCCAAGATTCCCTCACACCGATGGACACTTTTCATCCTTCTCCACTGATGGGAGGT 432  
Qy PheGlyLeuLysAsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520  
Db 431 TTGGACTCAAACACCCCGCTCCTCAGATCATGATCAAAAACACTCCCGTTCCAGCCAAT 372  
Qy ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540  
Db 371 CCTCCCAAACTTCAGTTCTGCCAAGTTTGCTTCTTTTCATCACACAGTATTCACGGGA 312  
Qy GlnValSerValGluIleGluTrrPgluLeuGlnLysGluAsnSerLysArgTrrPAsnPro 560  
Db 311 CAGGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAGAACAGCAACCGTGGAAATCCC 252  
Qy GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580  
Db 251 GAAATTCAGTACACTTCCAACATACAACAAGTCTGTTAATGTGGACTTTACTGTGGACACT 192  
Qy AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599  
Db 191 AATGGTGTATTTCAGAGCCTCGCCCATTCAGCACCAGATACCTGACTCGTAATCTG 135

RESULT 6

US-10-205-942-5  
; Sequence 5, Application US/10205942  
; Publication No. US20030053990A1  
; GENERAL INFORMATION:  
; APPLICANT: University of No. US20030053990A1th Carolina-Chapel Hill  
; APPLICANT: Rabinowitz, Joseph E.  
; APPLICANT: Samuleki, Richard J.  
; APPLICANT: Xiao, Weidong  
; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING THE SAME  
; FILE REFERENCE: 5470-186  
; CURRENT APPLICATION NUMBER: US/10/205,942  
; CURRENT FILING DATE: 2002-07-26  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 5  
; LENGTH: 8179  
; TYPE: DNA  
; ORGANISM: Adeno-associated virus  
US-10-205-942-5

Alignment Scores:  
Pred. No.: 5.19e-295 Length: 8179  
Score: 2801.00 Matches: 509  
Percent Similarity: 90.33% Conservative: 33  
Best Local Similarity: 84.83% Mismatches: 56  
Query Match: 86.16% Indels: 2  
DB: 15 Gaps: 2

US-09-807-802A-15 (1-599) x US-10-205-942-5 (1-8179)

Qy 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20  
Db 2544 ACGGCTCCGGAAAAAAGAGGCCGCTAGAGCACTCTCCTGTGGAGCCAGACTCCTCTCTCG 2603  
Qy 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40  
Db 2604 GGAACCGGAAAGCGGGCCAGCAGCCTGCAAGAAAAAGATTGAATTTGGTCAGACTGGA 2663

Qy 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60  
Db 2664 GACGCAGACTCAGTACCTGACCCCCCAGCCTCTCGGACAGCCACCAGCAGCCCTCTGGT 2723  
Qy ValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGly 80  
Db 2724 CTGGGAACATAACGATGGCTACAGGCAGTGGCGCACCAATGGCAGACATAACGAGGGC 2783  
Qy AlaAspGlyValGlyAsnAlaSerGlyAsnTrrPHisCysAspSerThrTrrPLeuGlyAsp 100  
Db 2784 GCCGACGGAGTGGTAATTCTCCGGAATGGCATTCGATTCCACATGGATGGGGCAG 2843  
Qy ArgValIleThrThrSerThrArgThrTrrPAlaLeuProThrTyrAsnAsnHisLeuTyr 120  
Db 2844 AGAGTCATCACCAACAGCACCCGAAACCTGGGCGCTGCCACCTACAAACACCACTCTAC 2903  
Qy LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140  
Db 2904 AAACAAATTTCCAGCCAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGC 2960  
Qy ThrProTrrPglyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrrP 160  
Db 2961 ACCCTTTGGGGGTATTTGACTTCAACAGATTCCACTGCCACTTTTCCACCAGCTGACTGG 3020  
Qy GlnArgLeuIleAsnAsnAsnTrrPglyPheArgProLysArgLeuAsnPhelysLeuPhe 180  
Db 3021 CAAAGACTCATCAACAACACTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTT 3080  
Qy AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200  
Db 3081 AACATTCAAGTCAAAGAGGTCAACGAGATGACGGTACGACGAGATTGCCAATAACCTT 3140  
Qy ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220  
Db 3141 ACCAGCACGGTTACGTGTTTACTGACTCGGAGTACAGCTCCCGTACGTGCTCGGGTCG 3200  
Qy AlaHisGlnGlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGly 240  
Db 3201 GCGCACCAAGGCTGTCTCCCGCGTTTCCAGCGGACGTCTTTCATGGTCCCTCAGTATGA 3260  
Qy TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260  
Db 3261 TACCTCACCTGAACAACGGAAGTCAAGCGGTGGAGCGCTCATCTTTTACTGCGCTGGAG 3320  
Qy TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280  
Db 3321 TACTTCCCTTCGCAGATGCTAAGGACTGGAATAACTTCCAATTCAAGCTATACCTTCGAG 3380  
Qy GluValPropheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300  
Db 3381 GATGTACCTTTTTCACAGCAGTACGCTCACAGCCAGAGTTTGGATCGCTTGATGAATCCT 3440  
Qy LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGln--AsnGlnSerGlySerAla 319  
Db 3441 CTTATTGATCAGTATCTGTACTACCTGAACAGAACGCAAGGAACAACCTCTGGAACAACC 3500  
Qy GlnAsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLys 339  
Db 3501 AACCAATCACGGCTGCTTTTAGCCAGGCTGGCGCTCAGTCTATGTCTTTGCAGGCGCAGA 3560  
Qy AsnTrrPLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsn 359  
Db 3561 AATTGGCTACCTGGGCGCTGTACCCGGCAACAGAGACTTTCAAAGACTGTCAACGACAAC 3620  
Qy AsnAsnSerAsnPheThrTrrPThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSer 379  
Db 3621 AACAAACAGTAACCTTTCTTGGACAGCGCGCCAGCAATAATATCATCTCAATGGCCGCGACTCG 3680  
Qy IleIleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePhePro 399  
Db 3681 CTGGTGAATCCAGGACAGCTATGGCCAGTCAAGAGACGATGAAGAAAAATTTTCCCT 3740  
Qy MetSerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAsp 419

Db 3741 ATGCACGGCAATCTAATATTTGGCAAGAGGACAAACGGCAAGTAACGCAGAATTAGAT 3800  
Qy 420 AsnValMetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArg 439  
Db 3801 AATGTAATGATTACGGATGAAGAAGAGATTTCGTACCACCAATCCTGTGGCAACAGAGCAG 3860  
Qy 440 PheGlyThrValAlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspVal 459  
Db 3861 TATGGAACCTGTGCAAAATAACTTGCAGAGCTCAAATACAGCTCCACGACTGGAACTGTC 3920  
Qy 460 HisAlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGly 479  
Db 3921 AATCATCAGGGGCGCTTACCTGGCATGGTGGCAAGATCGTGACGTGTACCTTCAAGGA 3980  
Qy 480 ProIleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGly 499  
Db 3981 CCTATCTGGGCAAGATTCTCTCACACGGATGGACACTTTCATCCTTCTCTCTGATGGGA 4040  
Qy 500 GlyPheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAla 519  
Db 4041 GGCTTTGGACTGAACATCCGCTCTCTCAATCATGATCAAAATACTCCGGTACCTGCG 4100  
Qy 520 AsnProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThr 539  
Db 4101 AATCCTTCGACCACCTTCAGTGGCGGCAAGTTTGTCTCTTCATCACACAGTACTCCACG 4160  
Qy 540 GlyGlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTyrAsn 559  
Db 4161 GGACAGGTCAGCGTGGAGATCGAGTGGGAGCTGAGAGGAAACAGCAACGCTGGAAT 4220  
Qy 560 ProGluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAsp 579  
Db 4221 CCCGAAATTCAGTACACTTCCAACTACAACAAGTCTGTTAATCGTGGACTTACCGTGGAT 4280  
Qy 580 AsnAsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599  
Db 4281 ACTAATGGCGTGATTTCAGAGCCTCGCCCCATTGGCACCAAGATACCTGACTCGTAATCTG 4340

RESULT 7  
US-09-804-898-1  
; Sequence 1, Application US/09804898  
; Patent No. US20020045264A1  
; GENERAL INFORMATION:  
; APPLICANT: DURING, MATTHEW  
; APPLICANT: XIAO, WEIDONG  
; TITLE OF INVENTION: PRODUCTION OF CHIMERIC CAPSID VECTORS  
; FILE REFERENCE: 102182-14  
; CURRENT APPLICATION NUMBER: US/09/804,898  
; CURRENT FILING DATE: 2001-03-13  
; PRIOR APPLICATION NUMBER: 60/189,110  
; PRIOR FILING DATE: 2000-03-14  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 4679  
; TYPE: DNA  
; ORGANISM: adeno-associated virus 2  
US-09-804-898-1

Alignment Scores:  
Pred. No.: 2.15e-291 Length: 4679  
Score: 2764.50 Matches: 496  
Percent Similarity: 89.98% Conservative: 43  
Best Local Similarity: 82.80% Mismatches: 59  
Query Match: 85.04% Indels: 1  
DB: 9 Gaps: 1

US-09-807-802A-15 (1-599) x US-09-804-898-1 (1-4679)

Qy 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20  
Db 2614 ACGGCTCCGGGAAAAAGAGGCGGTAGAGCACTCTCCTGTGGAGCCAGACTCCTCCTCG 2673

Qy 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40  
Db 2674 GGAACCGAAAGGCGGCAGCAGCCTGCAAGAAAAAGATTGAATTTTGGTCAGACTGGA 2733  
Qy 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60  
Db 2734 GACGCAGACTCAGTACCTGACCCCGCAGCCTCTCGGACAGCCACAGCAGCCCCCTCTGGT 2793  
Qy 61 ValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnGluGly 80  
Db 2794 CTGGGAACCTAATACGATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGCG 2853  
Qy 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100  
Db 2854 GCCGACGAGTGGGTAATTCCTCGGGAATTTGGCATTGGCATTCACATGGATGGGCGAC 2913  
Qy 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120  
Db 2914 AGAGTCATCACACCAGCACCCGAAACCTGGGCCCTGCCACCTACAACAACCACCTCTAC 2973  
Qy 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140  
Db 2974 AAACAAATTTCCAGCCAAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGC 3030  
Qy 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160  
Db 3031 ACCCCTGGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCAACCAGTGACTGG 3090  
Qy 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180  
Db 3091 CAAAGACTCATCAACAACAACCTGGGGATTCCGACCCCAAGAGACTCAACTTCNAGCTCTTT 3150  
Qy 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrIleAlaAsnAsnLeu 200  
Db 3151 AACATTCAAGTCAAAGAGGTCAAGCAGAAATGACGGTACGACGACGATTGCCAATAACCTT 3210  
Qy 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220  
Db 3211 ACCAGCAGGTTTCAGGTGTTTACTGACTCGGAGTACCAGTCCCGTACGTCTCTCGGCTCG 3270  
Qy 221 AlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGly 240  
Db 3271 GCGCATCAAGGATGCCTCCCGCGTTCACAGCAGACGTCTTCATGGTGCCACAGTATGGA 3330  
Qy 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260  
Db 3331 TACCTCACCTGAAACAACGGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCTGGAG 3390  
Qy 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280  
Db 3391 TACTTTCTCTCTCAGATGCTGCGTACCGGAAACAACACTTTTACCTTCAGCTACACTTTTGAG 3450  
Qy 281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300  
Db 3451 GACGTTCTTTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCT 3510  
Qy 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320  
Db 3511 CTCATCGACCAGTACCTGTATTACTTGAGCAGACAACAACACTCCCAAGTGGAACCCACG 3570  
Qy 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340  
Db 3571 CAGTCAAGGCTTCAGTTTCTACGCCGAGCGAGTGACATTCCGGACCAGTCTAGGAAC 3630  
Qy 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360  
Db 3631 TGGCTTCTCTGGACCTGTATTACCGCCAGCAGCGAGTATCAAAGACATCTGCGGATAACAAC 3690  
Qy 361 AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380  
Db 3691 AACAGTGAATACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTG 3750



QY 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400  
:::|||||  
Db 3751 GTGAATCCGGCCCGCATGGCAAGCCACAAAGGACGATGAAGAAAGTTTTTCTCTCAG 3810  
QY 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420  
|||||:::|||||:::|:::|:::|  
Db 3811 AGCGGGTTCTCATCTTTGGGAAGCAAGGCTCAGAGAAAACAAATGTGGACATTGAAAG 3870  
QY 421 ValMetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440  
|||||:::|||||:::|:::|:::|:::|:::|  
Db 3871 GTCATGATTACAGACGAAGAGGAAATCAGGACCAACCAATCCCGTGGCTACGAGCAGTAT 3930  
QY 441 GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis 460  
|||||:::|||||:::|:::|:::|:::|:::|  
Db 3931 GGTCTGTATCTACCAACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAAC 3990  
QY 461 AlaMetGlyAlaLeuProGlyMetValTyrGlnAspArgAspValTyrLeuGlnGlyPro 480  
|||:::|||||:::|:::|:::|:::|:::|  
Db 3991 ACACAAGCGTTCTTCCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCC 4050  
QY 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500  
|||||:::|||||:::|:::|:::|:::|:::|  
Db 4051 ATCTGGCAAGATTCCACACACGACGGACGACATTTTACCCCTCTCCCTCATGGGTGA 4110  
QY 501 PheGlyLeuLysAsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520  
|||||:::|||||:::|:::|:::|:::|:::|  
Db 4111 TTCGGACTTAAACACCCCTCTCCACAGATTCTCATCAAGAAACACCCCGGTACTCCGAAT 4170  
QY 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540  
|||:::|||||:::|:::|:::|:::|:::|  
Db 4171 CCTTCGACCACCTTCAGTGGGCAAGTTTGTCTTCTCATCACACAGTACTCCACGGGA 4230  
QY 541 GlnValSerValGluIleGluTyrGluLeuGlnLysGluAsnSerLysArgTrpAsnPro 560  
|||:::|||||:::|:::|:::|:::|:::|  
Db 4231 CAGGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAGGAAACAGCAACCGCTGGAATCCC 4290  
QY 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580  
|||:::|||||:::|:::|:::|:::|:::|  
Db 4291 GAAATTCAGTACACTTCCAACTACAACAAGTCTGTTAATGTGGACTTTACTGTGGACACT 4350  
QY 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599  
|||||:::|||||:::|:::|:::|:::|:::|  
Db 4351 AATGGCGTGTATTCAGAGCGCTGCCCCATTGGCACCAGATACCTGACTCGTAATCTG 4407

RESULT 8

US-09-945-681-10  
; Sequence 10, Application US/09945681  
; Patent No. US20020064878A1  
; GENERAL INFORMATION:  
; APPLICANT: UNIVERSITE DE NANTES  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR RECOMBINANT  
; TITLE OF INVENTION: ADENO-ASSOCIATED VIRUS PRODUCTION  
; FILE REFERENCE: B4182AA - UNIVERSITE DE NANTES  
; CURRENT APPLICATION NUMBER: US/09/945,681  
; CURRENT FILING DATE: 2001-09-05  
; PRIOR APPLICATION NUMBER: PCT/EP 00/01854  
; PRIOR FILING DATE: 2000-03-05  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 4679  
; TYPE: DNA  
; ORGANISM: adeno-associated virus 2  
US-09-945-681-10

Alignment Scores:  
Pred. No.: 2.15e-291 Length: 4679  
Score: 2764.50 Matches: 496  
Percent Similarity: 89.98% Conservative: 43  
Best Local Similarity: 82.80% Mismatches: 59  
Query Match: 85.04% Indels: 1  
DB: 9 Gaps: 1

US-09-807-802A-15 (1-599) x US-09-945-681-10 (1-4679)  
QY 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20  
|||||:::|||||:::|:::|:::|:::|:::|  
Db 2614 ACGGCTCCGGGAAAAAAGAGGCCCGGTAGAGCACTCTCCTGTGGAGCCAGACTCCTCCTCG 2673  
QY 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40  
|||:::|||||:::|:::|:::|:::|:::|  
Db 2674 GGAACCGGAAAGCGCGGCCAGCAGCTGCAAGAAAAGATTGAATTTGGTCAGACTGGA 2733  
QY 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60  
|||||:::|||||:::|:::|:::|:::|:::|  
Db 2734 GACGCAGACTCAGTACCTGACCCCGAGCTCTCGGACAGCCACCGACCCCTCTCTGGT 2793  
QY 61 ValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGly 80  
:::|||||:::|:::|:::|:::|:::|  
Db 2794 CTGGGAACATAATACGATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGGC 2853  
QY 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100  
|||||:::|||||:::|:::|:::|:::|:::|  
Db 2854 GCCGACGGAGTGGTAATTCTCTCGGAAATTTGGCATTTCCATTCACATGGATGGCGAC 2913  
QY 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120  
|||||:::|||||:::|:::|:::|:::|:::|  
Db 2914 AGAGTCATCACACCAGCACCCGAACTGGGCCCTGCCACCTTACCAACAACCCCTCTAC 2973  
QY 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140  
|||||:::|||||:::|:::|:::|:::|:::|  
Db 2974 AAACAAATTTCCAGCCAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGC 3030  
QY 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160  
|||||:::|||||:::|:::|:::|:::|:::|  
Db 3031 ACCCCTTGGGGTATTTTGACTTTCACAGATTCCACTGCCACTTTTCCACCGTACTGG 3090  
QY 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180  
|||||:::|||||:::|:::|:::|:::|:::|  
Db 3091 CAAAGACTCATCAACAACAACCTGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTT 3150  
QY 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200  
|||||:::|||||:::|:::|:::|:::|:::|  
Db 3151 AACATTCAAGTCAAAGAGGTACCGCAGAAATGACGGTACGACGACGATTGCCAATAACCTT 3210  
QY 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220  
|||||:::|||||:::|:::|:::|:::|:::|  
Db 3211 ACCAGCAGGTTCAGGTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTCCTCGGCTCG 3270  
QY 221 AlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGly 240  
|||||:::|||||:::|:::|:::|:::|:::|  
Db 3271 GCGCATCAAGGATGCCCTCCCGCGTTCCAGCAGACGCTCTTCATGTGTGCCACAGTATGA 3330  
QY 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260  
|||||:::|||||:::|:::|:::|:::|:::|  
Db 3331 TACCTCACCTGAACAACGGGAGTCAGGAGTAGGACGCTCTTCACTTTACTGCTGGAG 3390  
QY 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280  
|||||:::|||||:::|:::|:::|:::|:::|  
Db 3391 TACTTTCTCTCAGATGTCGTACCGTACCGGAAACAACCTTTACCTTCAGCTACACTTTTGAG 3450  
QY 281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300  
:::|||||:::|:::|:::|:::|:::|  
Db 3451 GACGTTCTTTCCACAGCAGCTACGCTCAAGCCAGCAGAGTCTGGACCGTCTCATGAATCCT 3510  
QY 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320  
|||||:::|||||:::|:::|:::|:::|:::|  
Db 3511 CTCATCGACCAGTACCTGTATTACTTTGAGCAGACAACAACACTCCCAAGTGGAAACCACCG 3570  
QY 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340  
|||:::|||||:::|:::|:::|:::|:::|  
Db 3571 CAGTCAAGGCTTCAGTTTCTCAGGCCGGAGCGAGTGACATTCGGGACCACTTAGGAAC 3630  
QY 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360  
|||||:::|||||:::|:::|:::|:::|:::|  
Db 3631 TGGCTTCTCGGACCCTGTTTACCGCCAGCAGCGAGTATCAAAGACATCTGCGGATACCAAC 3690

|   |           |  |      |
|---|-----------|--|------|
| Qy  | 361       | AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle   | 380  |
| Dd  | 3691      | AACAGTGAATACTCGTGGACTCGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTG  | 3750 |
| Qy  | 381       | IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet   | 400  |
| Dd  | 3751      | GTGAATCCGGCCCGCCATGGCAAGCCACAGGACGATGAAGAAAGTTTTCCTCAG         | 3810 |
| Qy  | 401       | SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn   | 420  |
| Dd  | 3811      | AGCGGGTTCTCATCTTTGGGAAGCAAGGCTCAGAGAAACAATGTGGACATTGGAAG       | 3870 |
| Qy  | 421       | ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe      | 440  |
| Dd  | 3871      | GTCAATGATACAGACGAAGAGGAATCAGGACCAACCAATCCCGTGGCTACGGAGCAGTAT   | 3930 |
| Qy  | 441       | GlyThrValAlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHis      | 460  |
| Dd  | 3931      | GGTTCTGTATCTACCAACCTCCAGAGAGGCAACAGACAAGCAGCTACCGAGATGTCAAC    | 3990 |
| Qy  | 461       | AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro   | 480  |
| Dd  | 3991      | ACACAAGGCGTTCTCCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCC     | 4050 |
| Qy  | 481       | IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly   | 500  |
| Dd  | 4051      | ATCTGGGCAAGATTCCACACACGGACGGACATTTTCACCCCTCTCCCTCATGGGTGA      | 4110 |
| Qy  | 501       | PheGlyLeuLysAsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsn   | 520  |
| Dd  | 4111      | TTCCGGACTTAAACACCCCTCCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCCAAT | 4170 |
| Qy  | 521       | ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly   | 540  |
| Dd  | 4171      | CCTTCGACCACCTTCAGTGGCGGAAAGTTGTCTTCTTCATCACACAGTACTCCACGGGA    | 4230 |
| Qy  | 541       | GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro   | 560  |
| Dd  | 4231      | CAGGTCAGCGTGAGATCGAGTGGGAGCTGCAGAAGGAAACAGCAAAACGCTGGAATCCC    | 4290 |
| Qy  | 561       | GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn   | 580  |
| Dd  | 4291      | GAAATTCAGTACACTTCCAACCTACAACAAGTCTGTTAATGTGGACTTTACTGTGGACACT  | 4350 |
| Qy  | 581       | AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu      | 599  |
| Dd  | 4351      | AATGGCGTGATTCAGAGCCTCGCCCCATTTGGCACCAGATACCTGACTCGTAATCTG      | 4407 |
| RESULT 9  |           |  |      |
| US-10-136-819-6   |           |  |      |
| ; Sequence 6, Application US/10136819   |           |  |      |
| ; Publication No. US20030166593A1   |           |  |      |
| ; GENERAL INFORMATION:  |           |  |      |
| ; APPLICANT: Chien, Kenneth   |           |  |      |
| ; APPLICANT: Hoshijima, Masahiko  |           |  |      |
| ; TITLE OF INVENTION: No. US20030166593A1-viral vesicle vector for cardiac specific gen |           |  |      |
| ; FILE REFERENCE: 6627-PA1198   |           |  |      |
| ; CURRENT APPLICATION NUMBER: US/10/136,819   |           |  |      |
| ; CURRENT FILING DATE: 2002-04-30   |           |  |      |
| ; PRIOR APPLICATION NUMBER: 60/287,423  |           |  |      |
| ; PRIOR FILING DATE: 2001-04-30   |           |  |      |
| ; NUMBER OF SEQ ID NOS: 18  |           |  |      |
| ; SOFTWARE: Patentin version 3.1  |           |  |      |
| ; SEQ ID NO 6   |           |  |      |
| ; LENGTH: 4679  |           |  |      |
| ; TYPE: DNA   |           |  |      |
| ; ORGANISM: adeno-associated virus 2  |           |  |      |
| US-10-136-819-6   |           |  |      |
| Alignment Scores:   |           |  |      |
| Pred. No.:  | 2,15e-291 | Length:  | 4679 |
| Score:  | 2764.50   | Matches:   | 496  |

|  |      |  |      |
|--|------|--|------|
| Percent Similarity: 89.98%                           |      | Conservative: 43   |      |
| Best Local Similarity: 82.80%                        |      | Mismatches: 59   |      |
| Query Match: 85.04%                                  |      | Indels: 1  |      |
| DB: 13   |      | Gaps: 1  |      |
| US-09-807-802A-15 (1-599) x US-10-136-819-6 (1-4679) |      |  |      |
| Qy   | 1    | ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer   | 20   |
| Dd   | 2614 | ACGGCTCCGGGAAAAAAGAGCCGCTAGAGCACTCTCCTGTGGAGCCAGACTCCTCCTCG    | 2673 |
| Qy   | 21   | GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly   | 40   |
| Dd   | 2674 | GGAACCGGAAAGGGCGGCCAGCAGCCTGCAAGAAAAAAGATTGAATTTTGGTCAGACTGGA  | 2733 |
| Qy   | 41   | AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla   | 60   |
| Dd   | 2734 | GACGCAGACTCAGTACTGACCCCGCCAGCCTCTCGACAGCCACCAGCAGCCCCCTCTGGT   | 2793 |
| Qy   | 61   | ValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGly   | 80   |
| Dd   | 2794 | CTGGGAACCTAATACGATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGC   | 2853 |
| Qy   | 81   | AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp   | 100  |
| Dd   | 2854 | GCCGACGGAGTGGTAATTCCTCGGGAAATTGGCATTCGATTCCACATGGATGGGCGAC     | 2913 |
| Qy   | 101  | ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr   | 120  |
| Dd   | 2914 | AGAGTCATCACCCACGACCCCGAACCTGGGCCCCCTGCCACCTACAACAACCCACCTCTAC  | 2973 |
| Qy   | 121  | LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer   | 140  |
| Dd   | 2974 | AAACAAATTTCCAGCCCAATCA--GGAGCCTCGAAGCACAATCACTACTTTGGCTACAGC   | 3030 |
| Qy   | 141  | ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp   | 160  |
| Dd   | 3031 | ACCCCTTGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCCACCACGTGACTGG   | 3090 |
| Qy   | 161  | GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe   | 180  |
| Dd   | 3091 | CAAAGACTCATCAACAACAACCTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTT | 3150 |
| Qy   | 181  | AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu   | 200  |
| Dd   | 3151 | AACATTCAAGTCAAGAGGTCAAGCAGAAATGACGGTACGACGAGATTGCCAATAACCTT    | 3210 |
| Qy   | 201  | ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer   | 220  |
| Dd   | 3211 | ACCAGCACGGTTTCAGGTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTCTCGGCTCG   | 3270 |
| Qy   | 221  | AlaHisGlnGlyCysLeuProPropheProAlaAspValPheMetIleProGlnTyrGly   | 240  |
| Dd   | 3271 | GCGCATCAAGGATGCCCTCCGCGCTTCCCAGCAGACGCTTTCATGGTGGCCACAGTATGGA  | 3330 |
| Qy   | 241  | TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu   | 260  |
| Dd   | 3331 | TACCTCACCTGAAACAACGGGAGTCAGGCAGTAGGACGCTCTTTCATTTTACTGCTGGAG   | 3390 |
| Qy   | 261  | TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu   | 280  |
| Dd   | 3391 | TACTTTCTCTCAGATGCTGCTGCTACCGGAAACAACCTTTACCTTCAGCTACACTTTTGAG  | 3450 |
| Qy   | 281  | GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro   | 300  |
| Dd   | 3451 | GACGTTCTCTTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCT   | 3510 |
| Qy   | 301  | LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln   | 320  |
| Dd   | 3511 | CTCATCGACCAGTACCTGTATTACTTTGAGCAGAAACAACACTCCAAGTGGAACCAACCG   | 3570 |
| Qy   | 321  | AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn   | 340  |

Db 3571 CAGTCAAGGCTTCAGTTTCTCAGGCCGGAGCGAGTGACATTCGGGACCAGTCTAGGAAC 3630

Qy 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360

Db 3631 TGGCTTCCTGGACCCCTGTTACCGCCAGCAGCGAGTATCAAAGACATCTCGGATAACAAC 3690

Qy 361 AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380

Db 3691 AACAGTGAATACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTG 3750

Qy 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400

Db 3751 GTGAATCCGGGCCCGCCCATGGCAAGCCACAAGGACGATGAAGAAAAAGTTTTCCTCAG 3810

Qy 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420

Db 3811 AGCGGGGTTCATCTTTGGGAAGCRAAGGCTCAGAGAAAAACAATGTGGACATTGAAAG 3870

Qy 421 ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440

Db 3871 GTCATGATTACAGACGAAGAGGAATCAGGACAACCAATCCCCTGGTACGGAGCAGTAT 3930

Qy 441 GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis 460

Db 3931 GGTTCTGTATCTACCAACCTCCAGAGGCAACAGACAAGCAGTACCGCAGATGTCAAC 3990

Qy 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 480

Db 3991 ACACAAGGCGTCTTCCAGGCATGGTTCGCGAGGACAGAGATGTGTACCTTCAGGGGCC 4050

Qy 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500

Db 4051 ATCTGGGCAAGATTCCACACACGACGACATTTTCACCCCTCCTCCCTCATGGGTGGA 4110

Qy 501 PheGlyLeuLysAsnProProGlnIleLeuLysAsnThrProValProAlaAsn 520

Db 4111 TTCGGACTTAACACCCCTCCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAAT 4170

Qy 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540

Db 4171 CCTTCGACCACCTTCAGTGGCGGCAAGTTTGCTTCCTTCATCAACAGTACTCCACGGGA 4230

Qy 541 GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro 560

Db 4231 CAGGTCAAGTGGAGATCGAGTGGGAGCTGCAGAAGGAAAAACAGCAACGCTGGAATCCC 4290

Qy 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580

Db 4291 GAAATTCAGTACACTTCCAACACTACAACAGTCTGTTAATGTGGACTTACTGTGGACACT 4350

Qy 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599

Db 4351 AATGGCGTGTATTACAGAGCCTCGCCCCATTGGCACCAGATACCTGTACTCGTAATCTG 4407

RESULT 10

US-10-038-972A-12

; Sequence 12, Application US/10038972A

; Publication No. US20020192823A1

GENERAL INFORMATION:

; APPLICANT: J. Bartlett

; TITLE OF INVENTION: AAV VECTORS AND METHODS

; FILE REFERENCE: 28335/36996US

; CURRENT APPLICATION NUMBER: US/10/038,972A

; CURRENT FILING DATE: 2002-01-04

; PRIOR APPLICATION NUMBER: US 60/260,124

; PRIOR FILING DATE: 2001-01-05

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 12

; TYPE: DNA

; ORGANISM: adeno-associated virus 2

US-10-038-972A-12

Alignment Scores:

Pred. No.: 2.15e-291 Length: 4679

Score: 2764.50 Matches: 496

Percent Similarity: 89.98% Conservative: 43

Best Local Similarity: 82.80% Mismatches: 59

Query Match: 85.04% Indels: 1

DB: 14 Gaps: 1

US-09-807-802A-15 (1-599) x US-10-038-972A-12 (1-4679)

Qy 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20

Db 2614 ACGGCTCCGGGAAAAAAGAGGCCGGTAGAGCACTCTCCTGTGGAGCCAGACTCCTCTCG 2673

Qy 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40

Db 2674 GGAACCGGAAGGGCGGCAGCAGCCTGCAAGAAAAAAGATTGAATTTTGGTCAGACTGGA 2733

Qy 41 AspSerGluSerValProAspProGlnProLeuGluProProAlaThrProAlaAla 60

Db 2734 GACGCAGACTCAGTACCTGACCCCGAGCCTCTCGGACAGCCACCAGCAGCCCTCTCGT 2793

Qy 61 ValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGly 80

Db 2794 CTGGGAACATAACGATGGCTACAGGCAGTGGCGCACTTCCGCAATGGCATTCGATTCACATGGATGGGCGAC 2853

Qy 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100

Db 2854 GCCGACGGAGTGGGTAATCTCTCGGAAATTGGCAATTCGATTCACATGGATGGGCGAC 2913

Qy 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120

Db 2914 AGAGTCATCACCAACAGCACCCGAACCTGGGCCCTGCCACCTACAACAACCCACTCTAC 2973

Qy 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140

Db 2974 AAACAAATTTCCAGCCCAATCA--GGAGCCTCGAACCAATCACTACTTTGGCTACAGC 3030

Qy 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160

Db 3031 ACCCTTGGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCACCACGTGACTGG 3090

Qy 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180

Db 3091 CAAAGACTCATCAACAACAACTGGGATTCGACCCCAAGAGACTCAACTTCAAGCTCTTT 3150

Qy 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200

Db 3151 AACATTCAAGTCAAAGAGGTACGCAGAAATGACGGTACGACGACGATTGCCAATAACCTT 3210

Qy 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220

Db 3211 ACCAGCAGGTTTCAAGTGTACTGACTCGGAGTACAGCTCCGCTCCGCTCGGCTCG 3270

Qy 221 AlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGly 240

Db 3271 GCGCATCAAGGATGCCCTCCGCGCTTCCAGCAGACGCTTTCATGGTGGCCACAGTATGGA 3330

Qy 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260

Db 3331 TACCTCACCCCTGAACAACGGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCCTGGAG 3390

Qy 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280

Db 3391 TACTTTCCTTCTCAGATGCTGGTACCGGAAACAACATTACCTTCAGCTACACATTTTGAG 3450

Qy 281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300

Db 3451 GACGTTCTTTCCACAGCAGTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCT 3510

Qy 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320



Db 3511 CTCATCGACCAGTACCTGTATTACTTGAGCAGAAACAACACTCCAAGTGAACCAACCACG 3570  
Qy 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340  
Db 3571 CAGTCAAGCTTCAGTTTCTCAGGCCGAGCGAGTGACATTCGGGACCACTCTAGGAAC 3630  
Qy 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360  
Db 3631 TGGCTTCTGGACCCCTGTTACCGCCAGCAGCGAGTATCAAGACATCTGCGGATAACAAC 3690  
Qy 361 AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380  
Db 3691 AACAGTGAATACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTG 3750  
Qy 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400  
Db 3751 GTGAATCCGGGCCCGCCATGGCAAGCCACAGGACGATGAAGAAAAGTTTTCCTCAG 3810  
Qy 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420  
Db 3811 AGCGGGTTCTCATCTTTGGGAAGCAAGCTCAGAGAAACAATGTGGACATTGAAAAG 3870  
Qy 421 ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440  
Db 3871 GTCATGATTACAGACGAAGAGAAATCAGACAACCAATCCCTGGTACGGAGCAGTAT 3930  
Qy 441 GlyThrValAlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHis 460  
Db 3931 GGTTCGTATCTACCAACCTCCAGAGAGGCAACAGACAAGCAGTACCGCAGATGTCAAC 3990  
Qy 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 480  
Db 3991 ACACAAGGCGTTCTTCCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCC 4050  
Qy 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500  
Db 4051 ATCTGGCAAGATTCCACACACGACGGACATTTTCACCCCTCTCCCTCATGGGTGA 4110  
Qy 501 PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520  
Db 4111 TTCGGACTTAAACACCCCTCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAAT 4170  
Qy 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540  
Db 4171 CCTTCGACCACCTTCAGTGGCGCAAGTTTGCTTCCTTCATCACACAGTACTCCACGGGA 4230  
Qy 541 GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro 560  
Db 4231 CAGGTCAAGTGGAGATCGAGTGGGAGCTGCAGAAGGAAACAGCAACGCTGGAATCCC 4290  
Qy 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580  
Db 4291 GAAATTCAGTACACTTCCAACACTACAACAGTCTGTAAATGTGGACTTTACTGTGGACACT 4350  
Qy 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599  
Db 4351 AATGGCGTGATTTCAGAGCCCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTG 4407

RESULT 11  
US-10-291-583-120  
; Sequence 120, Application US/10291583  
; Publication No. US20030138772A1  
; GENERAL INFORMATION:  
; APPLICANT: Gao, Guangping  
; APPLICANT: Wilson, James M.  
; APPLICANT: Alvira, Mauricio  
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus  
; FILE REFERENCE: UPN-02735USA  
; CURRENT APPLICATION NUMBER: US/10/291,583  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: US 60/350,607  
; PRIOR FILING DATE: 2001-11-13

; PRIOR APPLICATION NUMBER: US 60/341,117  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: US 60/377,066  
; PRIOR FILING DATE: 2002-05-01  
; PRIOR APPLICATION NUMBER: US 60/386,675  
; PRIOR FILING DATE: 2002-06-05  
; NUMBER OF SEQ ID NOS: 120  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 120  
; LENGTH: 2205  
; TYPE: DNA  
; ORGANISM: adeno-associated virus serotype, clone A3.1vp1  
US-10-291-583-120

Alignment Scores:  
Pred. No.: 1 42e-291 Length: 2205  
Score: 2761.50 Matches: 497  
Percent Similarity: 89.98% Conservative: 42  
Best Local Similarity: 82.97% Mismatches: 59  
Query Match: 84.94% Indels: 1  
DB: 13 Gaps: 1

US-09-807-802A-15 (1-599) x US-10-291-583-120 (1-2205)  
Qy 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20  
Db 412 ACGGCTCCTGGAAAAAAGAGACCTATAGAGCAGTCTCCTGCAGAACCGGACTCTCTCG 471  
Qy 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40  
Db 472 GGCATCGCAAAATCAGGCCAGCAGCCCCGCTAAGAAAAAGACTCAATTTTGGTCAGACTGGC 531  
Qy 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60  
Db 532 GACACAGAGTCAGTCCCGACCCCTCAACCAATCGGAGAACCCCCCGCAGCCCCCTCTGGT 591  
Qy 61 ValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGly 80  
Db 592 GTGGGATCTAATACAAATGCTTCAGGCGGTGGGGACCAATGGCAGACAATAACGAAGGC 651  
Qy 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100  
Db 652 GCCACGGAGTGGGTAACTCCTCGGAAATTTGGCAATTCGATTCACATGGATGGCGAC 711  
Qy 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120  
Db 712 AGAGTTATCACCCAGCAGCAAGAACCTGGGCCCTCCCCACCTACAATAATCACCTCTAC 771  
Qy 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140  
Db 772 AAGCAAAATCTCCAGCGAATCG--GGAGCCACCAACGACCACTACTTCGGCTACAGC 828  
Qy 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160  
Db 829 ACCCCCTGGGGGTATTTTGACTTTAACAGATTCCACTGTCACTTCTCACCACGTGACTGG 888  
Qy 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180  
Db 889 CAGCGACTCATCAACAACAACACTGGGGATTTAGACCCCAAGAACTCAATTTCAAGCTCTTC 948  
Qy 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrIleAlaAsnAsnLeu 200  
Db 949 AACATCCAAGTCAAGGAGGTACGAGAGTATGATGGAACCAACGACCATCGCCAATAACCTT 1008  
Qy 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220  
Db 1009 ACCAGCAGGTGCAGGTCTTTCACAGACTCTGAGTACCAGCTGCCCTACGTCTCGTTTCG 1068  
Qy 221 AlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGly 240  
Db 1069 GCTCACCAGGGCTGCCTTCCGCCGTTCCAGCAGACGCTTTCATGATTCCTCAGTACGGC 1128  
Qy 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260

```

Db 1129 TACTTGACTCTGAACAATGCGAGCCAAGCGGTAGGACGTTCTTCTATTCTACTGCTAGAG 1188
Qy 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280
Db 1189 TATTTTCCCTCTCAGATGCTGAGGACGGGAAACAACATTACCTTCAGCTACACTTTTGAA 1248
Qy 281 GluValPropHeHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300
Db 1249 GACGTGCCTTTCCACAGCAGCTACGCGCACAGCCAGAGTCTGGATCGGCTGATGAATCCT 1308
Qy 301 LeuIleAspGlnTyrLeuTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320
Db 1309 CTCATTGACCAGTACCTGTATTACCTTGAGCAAACTCAGGGTACAAGTGGAAACAACGCAG 1368
Qy 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340
Db 1369 CAATCGAGACTGCAGTTTCAGCCAAGCTGGGCCTAGCTCCATGGCTCAGCAGGCCAAAAAC 1428
Qy 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360
Db 1429 TGGCTACCGGGACCCAGCTACCGACAGCAGCGAATGTCTAAGACGGCTAATGACAACAAC 1488
Qy 361 AsnSerAsnPheThrTyrThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380
Db 1489 AACAGTGAATTTGCTTGGAGTGCAGCCACCACCAATATTACCTGAATGGAAGAATTTCTCTG 1548
Qy 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400
Db 1549 GTCAATCCCGGGCCCCCAATGGCCAGTCAACAGGACGATGAGGAAAAGTATTTCGCCATG 1608
Qy 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420
Db 1609 CACGGAATCTCATCTTTGGAAAAACAAGGCACAGGAACCTACCAATGTGGACATTGAATCA 1668
Qy 421 ValMetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440
Db 1669 GTGCTTATTACAGACGAAGAAGAAATCAGAACAACTAATCTGTGGCTACAGAACAATAC 1728
Qy 441 GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis 460
Db 1729 GGACAGGTTGCCACCACCAACCATCAGAGTCAGAACACCACAGCTTCCTATGGAAGTGGAC 1788
Qy 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 480
Db 1789 ACCCAGGGAATCTTACCTCGAATGGTGTGGCAGGACCGCGATGCTATCTTCAAGTCCC 1848
Qy 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500
Db 1849 ATTTGGGCCAAAACCTCCTCACACGGACGGACACTTTTCATCTTCTCCGCTCATGGGAGGC 1908
Qy 501 PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520
Db 1909 TTTGGACTGAACACACCCTCCTCCCCAGATCCTGTATCAAAAAACACACCTGTGCCAGCGAAT 1968
Qy 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540
Db 1969 CCGCGGACCACCTTTCACCTCCTGGAAAAGTTTGCTTCGTTTCATTACCCAGTATTCACCGGA 2028
Qy 541 GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro 560
Db 2029 CAGGTGAGCGTGGAAATAGAGTGGAGCTGCAGAAAGAAAACAGCAACCGCTGGAAACCCA 2088
Qy 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580
Db 2089 GAAATTGAGTACACCTCCAACTACAACAAGTCGGTGAATGTGGAGTTTACCCTGGACGCA 2148
Qy 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599
Db 2149 AACGGTGTATTCTGAACCCCGCCCTATTGGCACTCGTTACCTTACCCGGAACTTG 2205
```

RESULT 12

US-10-291-583-56

```

; Sequence 56, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56
; LENGTH: 3122
; TYPE: DNA
; ORGANISM: new AAV serotype, clone A3.7
US-10-291-583-56
```

```

Alignment Scores:
Pred. No.: 1.83e-290 Length: 3122
Score: 2753.50 Matches: 496
Percent Similarity: 89.82% Conservative: 42
Best Local Similarity: 82.80% Mismatches: 60
Query Match: 84.70% Indels: 1
DB: 13 Gaps: 1
```

US-09-807-802A-15 (1-599) x US-10-291-583-56 (1-3122)

```

Qy 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20
Db 1258 ACGGTCCTCGAAAAAAGAGACCTATAGAGCAGTCTCCTGCAGAACCGGACTCTTCCTCG 1317
Qy 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40
Db 1318 GGCATCGGCAAAATCAGGCCAGCCCGCTAAGAAAAGACTCAATTTTGGTCAGACTGGC 1377
Qy 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60
Db 1378 GACACAGAGTCAGTCCCAGACCCCTCAACCAATCGGAGAACCCCGCCGCTCTGGT 1437
Qy 61 ValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGly 80
Db 1438 GTGGGATCTAATACAAATGGCTTCAGGCGGTGGGCGCACCAATGGCAGACAATAACGAAGGC 1497
Qy 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100
Db 1498 GCCGACGGAGTGGGTAATTCTCGGGAATTTGGCATTTCCACATGGATGGGCGGAC 1557
Qy 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120
Db 1558 AGAGTTATCACCACCAGCAACAAGAACCTGGGCGCTCCCACTACATAATAATCGCCTCTAC 1617
Qy 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140
Db 1618 AAGCAAATCTCCAGCGAATCG--GGAGCCCAACAGACAACCACTACTTCGGGTACAGC 1674
Qy 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160
Db 1675 ACCCCCTGGGGGTATTTTGACTTTTAAACAGATTCCACTGTCTCACTTCTCACCACGTGACTGG 1734
Qy 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180
Db 1735 CAGCGACTCATCAACAACAACACTGGGGATTTAGACCCCAAGAACTCAATTTCAAGCTCTTC 1794
```

Qy 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200  
Db 1795 AACATCCAAGTCAAGGAGGTACGCAGAATGATGGAACCAACGACCATCGCCAATAACCTT 1854  
Qy 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220  
Db 1855 ACCAGCAGGTGCAGGTCTTCACAGACTCTGAGTACCAGTGCCTACGTCCTCGGTTGG 1914  
Qy 221 AlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGly 240  
Db 1915 GCTCACAGGGCTGCCCTTCGCCCGTTCAGCAGACAGCTCTTCATGATTCCTCAGTACGGC 1974  
Qy 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260  
Db 1975 TACTTGACTCTGAAACAATGGCAGCCAGCGGTAGGACGTTCTTCATTCTACTGTCTAGAG 2034  
Qy 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280  
Db 2035 TATTTTCCCTCTCAGATGCTGAGGACGGGAAACAACCTTCACCTCAGCTACACCTTTTGAA 2094  
Qy 281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300  
Db 2095 GACGTGCCCTTCCACAGCAGCTACGCGCACAGCCAGAGTCTGGATCGGCTGATGAATCCT 2154  
Qy 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320  
Db 2155 CTCATTGACCAGTACCTGTATTACCTGAGCAAACTCAGGGTACAAGTGAACAACGCGAG 2214  
Qy 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340  
Db 2215 CAATCGAGACTGCAGTTCAGCCAAGCTGGGCTAGCTCCATGGCTCAGCAGGCCAAAAC 2274  
Qy 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360  
Db 2275 TGGCTACCGGGACCCAGCTACCGACAGCAGCGAATGTCTAAGACGGCTAATGACAACAAC 2334  
Qy 361 AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380  
Db 2335 AACAGTGAATTTGCTTGGACTGCAGCCACCAAAATATTACCTGAATGGAAGAAATTCCTG 2394  
Qy 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400  
Db 2395 GTCAATCCCGGGCCCCCAATGGCCAGTCAACAGGACGATGAGGAAAAGTATTTCCCCCATG 2454  
Qy 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420  
Db 2455 CACGGAAATCTCATCTTTGGAAAACAAGGCACAGGAACCTACCAATGTGACATTGAATCA 2514  
Qy 421 ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440  
Db 2515 GTGCTTATTACAGACGAAGAAGAAATCAGAACAACTAATCCTGTGGTACAGAACAAATAC 2574  
Qy 441 GlyThrValAlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHis 460  
Db 2575 GGACAGGTTGCCACCAACCATCAGAGTCAGAACACCACAGCTTCCTATGTGAAGTGTGGAC 2634  
Qy 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 480  
Db 2635 AGCCAGGGAATCTTACCTGGAATGGTGTGGCAGGACCGCGATGTCTATCTTCAAGGTCCC 2694  
Qy 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500  
Db 2695 ATTTGGGCCAAAACCTCTCACACGGACGGACACTTTTCATCTTCTCCGCTCATGGGAGGC 2754  
Qy 501 PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520  
Db 2755 TTTGGACTGAAACACCTCTCTCCAGATCCTGATCAAAACACACCTGTGCCAGCGAAT 2814  
Qy 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540  
Db 2815 CCCGGACCACTTTCACTCTCTGGAAAGTTTGTGTTCTATTACCCAGTATTCCACCGGA 2874  
Qy 541 GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro 560

Db 2875 CAGGTCAGCGTGGAATAGAGTGGGAGCTGCAGAAAGAAACAGCAACCGCTGGAACCCA 2934  
Qy 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580  
Db 2935 GAAATTCAGTACACCTCCAACACTACAACAAGTCGGTGAATGTGGAGTTTACCCGTGGACGCA 2994  
Qy 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599  
Db 2995 AACGGTGTTTATTCTGAACCCCGCCCTATTGGCACTCGTTACCTTACCCGGAACCTTG 3051  
RESULT 13  
US-10-291-583-54  
; Sequence 54, Application US/10291583  
; Publication No. US20030138772A1  
; GENERAL INFORMATION:  
; APPLICANT: Gao, Guangping  
; APPLICANT: Wilson, James M.  
; APPLICANT: Alvira, Mauricio  
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus  
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1el Sequences Identif.  
; FILE REFERENCE: UPN-02735USA  
; CURRENT APPLICATION NUMBER: US/10/291,583  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: US 60/350,607  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/341,117  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: US 60/377,066  
; PRIOR FILING DATE: 2002-05-01  
; PRIOR APPLICATION NUMBER: US 60/386,675  
; PRIOR FILING DATE: 2002-06-05  
; NUMBER OF SEQ ID NOS: 120  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 54  
; LENGTH: 3123  
; TYPE: DNA  
; ORGANISM: new AAV serotype, clone A3.4  
US-10-291-583-54  
Alignment Scores:  
Pred. No.: 3,03e-290 Length: 3123  
Score: 2751.50 Matches: 495  
Percent Similarity: 89.82% Conservative: 43  
Best Local Similarity: 82.64% Mismatches: 60  
Query Match: 84.64% Indels: 1  
DB: 13  
US-09-807-802A-15 (1-599) x US-10-291-583-54 (1-3123)  
Qy 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20  
Db 1249 ACGGTCCTCGGAAAAAGAGACCTATAGAGCAGTCTCCTGCAGAACCGGACTCTTCCTCG 1308  
Qy 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40  
Db 1309 GGCATCGCGGAATCAGGCCACGACCCCGCTAAGAAAAGACTCAATTTTGGTCAGACTGGC 1368  
Qy 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60  
Db 1369 GACACAGAGTCAGTCCAGACCCCTCAACCAATCGGAGAACCCCGCAGCCCCCTCTGGT 1428  
Qy 61 ValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGly 80  
Db 1429 GTGGATCTAATAACATGGCTTCAGGCGGTGGGGCACCAATGGCAGACGATAACGAAGGC 1488  
Qy 81 AlaAspGlyValGlyAsnAlaSerGlyGlyAlaProHisCysAspSerThrTrpLeuGlyAsp 100  
Db 1489 GCCGACGGAGTGGGTAATTCCTCGGAAAATTGGCATTTGCCATTCACATGGATGGCGAC 1548  
Qy 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120  
Db 1549 AGAGTTATCACCACCAGCACACGAACCTGGGCCCTCCCCACCTACAAATAATCACCTCTAC 1608



Qy 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140  
Db 1609 AAGCAAATCTCCAGCGAATCG--GGAGCCACCAACGACAAACCACACTACTTCGGGTACAGC 1665  
Qy 141 ThrProTyrGlyTyrPheAppPheAsnArgPheHisCysHisPheSerProArgAspTyr 160  
Db 1666 ACCCCCTGGGGGTATTTTGACTTTAACAGATTCCCACTGTCTCACTTCTCACCCACGTGACTGG 1725  
Qy 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180  
Db 1726 CAGCGACTCATCAACAACAACTGGGGATTTAGACCCCAAGAAACTCAATTTCAAGCTCTTC 1785  
Qy 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200  
Db 1786 AACATCCAAGTCAAGGAGGTACGCAGAGATGATGGAAACCACGACCATCGCCCAATAACCTT 1845  
Qy 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220  
Db 1846 ACCAGCACGGTGCAGGTCTTCACAGACTCTGAGTACCAGTGCCTACGTCCCTCGGTTCTG 1905  
Qy 221 AlaHisGlnGlyCysLeuProPropPheProAlaAspValPheMetIleProGlnTyrGly 240  
Db 1906 GCTCACCAGGGCTGCCTTCGCGCGTTCACAGAGTCTCCAGCAGACGTCTTCATGATTCCTCAGTACGGC 1965  
Qy 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260  
Db 1966 TACTTGACTCTGAACAATGGCAGCCAAGCGGTAGGACGTCTTTCATTCTACTGTCTAGAG 2025  
Qy 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280  
Db 2026 TATTTCCCTCTCAGATGCTGAGGACGGGAAACAACATTCACCTTCAGCTACACTTTTGAA 2085  
Qy 281 GluValPropPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300  
Db 2086 GACGTGCCTTTCCACAGCAGCTACCGCGCAGACCCAGAGTCTGGATCGGCTGATGAATCCT 2145  
Qy 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320  
Db 2146 CTCATTGACCAGTACCTGTATTACCTGAGCAAAACTCAGGTACAAGTGGAAACAACGACG 2205  
Qy 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340  
Db 2206 CAATCGAGACTGCAGTTTCAGCCAAGCTGGGCCTAGCTCCATGGCTCAGCAGGCCCAAAAC 2265  
Qy 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360  
Db 2266 TGGCTACCGGGACCCAGCTACCGACAGCAGCGAATGTCTAAGACGGCTAATGACAACAAC 2325  
Qy 361 AsnSerAsnPheThrTyrThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380  
Db 2326 AACAGTGAATTTGCTTGGACTGCGAGCCACCAATAATATTACCTGAATGGAGAAATTCCTGTG 2385  
Qy 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400  
Db 2386 GTCAATCCCGGGCCCCCAATGGCCAGTCACAAGGACGATGAGGAAAAGTATTTCCCCATG 2445  
Qy 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420  
Db 2446 CACGGAATCTCATCTTTGGAAAAACAAGGCACAGGAACCTACCAATGTGGACATTGAATCA 2505  
Qy 421 ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440  
Db 2506 GTGCTTATTACAGACGAAGAAGAAATCAGAACAACATAATCCTGTGGCTACAGAACAATAC 2565  
Qy 441 GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis 460  
Db 2566 GGACAGGTTGCCACCACCATCAGAGTCAGGACACCACAGCTTCCTATGGAAGTGTGAC 2625  
Qy 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 480  
Db 2626 AGCCAGGGAATCTTACCTGGAATGGTGTGGCAGGACCGCGATGTCTATCTTCAAGGTCCC 2685

Qy 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500  
Db 2686 ATTTGGGCCAAAACCTCCTCACACGGACGGACACTTTTCATCCTTCTCCGTCATGGAGGC 2745  
Qy 501 PheGlyLeuLysAsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520  
Db 2746 TTTGGACTGAAACACCCCTCCTCCCCAGATCCTGATCAAAAACACACCTGTGCCAGCGAAT 2805  
Qy 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540  
Db 2806 CCGCGACCACTTCACTCCTGGAAAGTTTGTCTCGTTCAATACCCAGTATTCACCCGGA 2865  
Qy 541 GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro 560  
Db 2866 CAGGTCAAGCTGGAAAATAGAGTGGGAGCTGCAGAAAAGAAAACAGCAAAACGCTGGAAACCCA 2925  
Qy 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580  
Db 2926 GAAATTCAGTACACTCCAACCTACAACAAGTCGGTGAATGTGGAGTTTACCGTGGACGCA 2985  
Qy 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599  
Db 2986 AACGGTGTATTATCTGAACCCCGCCCTATTGGCACTCGTTACCTTACCCCGGAACCTTG 3042

RESULT 14

US-10-291-583-55  
; Sequence 55, Application US/10291583  
; Publication No. US20030138772A1  
; GENERAL INFORMATION:  
; APPLICANT: Gao, Guangping  
; APPLICANT: Wilson, James M.  
; APPLICANT: Alvira, Mauricio  
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus  
; FILE REFERENCE: UPN-02735USA  
; CURRENT APPLICATION NUMBER: US/10/291,583  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: US 60/350,607  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/341,117  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: US 60/377,066  
; PRIOR APPLICATION NUMBER: US 60/386,675  
; PRIOR FILING DATE: 2002-06-05  
; NUMBER OF SEQ ID NOS: 120  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 55  
; LENGTH: 3113  
; TYPE: DNA  
; ORGANISM: new AAV serotype, clone A3.5  
US-10-291-583-55

Alignment Scores:  
Pred. No.: 5e-290 Length: 3113  
Score: 2749.50 Matches: 495  
Percent Similarity: 89.82% Conservative: 43  
Best Local Similarity: 82.64% Mismatches: 60  
Query Match: 84.57% Indels: 1  
DB: 13 Gaps: 1

US-09-807-802A-15 (1-599) x US-10-291-583-55 (1-3113)

Qy 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20  
Db 1249 ACGGCTCCTGGAAAAAAGAGACCTATAGAGCAGTCTCCTGCAGAACCGGACTCTTCCTCG 1308  
Qy 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40  
Db 1309 GGCATCGGCAAAATCAGCCAGCAGCCCCGCTAAGAAAAGACTCAATTTTGGTCAGACTGGC 1368  
Qy 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60

Db 1369 GACACAGAGTCAGTCCCGAGACCCCTCAACCAATCGGAGAACCCCGCAGCCCCCTTGGT 1428  
 Qy 61 ValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGly 80  
 Db 1429 GTGGATCTAATAACATGGCTTCAGCGGTGGGGCCACCATTGGCAGACAATAACGAGGC 1488  
 Qy 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100  
 Db 1489 GCCGACGAGTGGGTAAATTCCTCGGAAATTGGCATTTCCATGGATGGGCGAC 1548  
 Qy 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120  
 Db 1549 AGAGTTATCACACCAGCACACAGAACCTGGGCCCTCCCACTTACAATAATCACCTTAC 1608  
 Qy 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140  
 Db 1609 AAGCAAAATCTCCAGCGAATCG---GGAGCCACCAACGACACCACTACTTCGGCTACAGC 1665  
 Qy 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160  
 Db 1666 ACCCCCTGGGGTATTTTGACTTTAACAGATTCCACTGTCACTTCTCACCCAGCTGACTGG 1725  
 Qy 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180  
 Db 1726 CAGCGACTCATCAATAACAACCTGGGGATTAGACCCCAAGAACTCAATTTCAAGCTTTC 1785  
 Qy 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200  
 Db 1786 AACATCCAAGTCAAGGAGGTCAAGCAGAAATGATGGAACCAAGCACTGCGCAATAACCTT 1845  
 Qy 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220  
 Db 1846 ACCAGCACGGTGCAGGTCTTTCAGACTCTGAGTACCAGTGCCTTACGTCTCTCGGTTCC 1905  
 Qy 221 AlaHisGlnGlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGly 240  
 Db 1906 GCTCACCAAGGTGCTTCCGCCGTTCCAGCAGACAGTCTTTCATGATTCTTCAGTACGGC 1965  
 Qy 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260  
 Db 1966 TACTTGACTCTGAACAATGGCAGCCCAAGCGGTAGGACGTTCTTCATTCTACTGTCTAGAG 2025  
 Qy 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280  
 Db 2026 TATTTTCCCTCTCAGATGCTGAGGACGGGAACAACCTTCACCTTCAGCTACACTTTTGAA 2085  
 Qy 281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300  
 Db 2086 GACGTGCCTTTCCACAGCAGCTACGGGCACAGCCAGAGTCTGGATCGGTGATGAATCCT 2145  
 Qy 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320  
 Db 2146 CTCATTGACCAGTACCTGTATTACCTGAGCAAAACTCAGGTACAAGTGGAAACAACGCAG 2205  
 Qy 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340  
 Db 2206 CAATCGAGACTGCAGTTCAAACCAAGCTGGGCCTAGCTCCATGGCTCAGCAGGCCAAAAAC 2265  
 Qy 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360  
 Db 2266 TGGCTACCGGGACCCAGCTACCGACAGCAGCGAATGTCTAAGACGGCTAATGACAACAAC 2325  
 Qy 361 AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380  
 Db 2326 AACAGTGAATTTGCTTGGACTGCAGCCACCAAAATATTACCGAATGGAAGAAATTTCTCTG 2385  
 Qy 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400  
 Db 2386 GTCAATCCCGGGCCCCCAATGGCCAGTCAACAAGGACGATGAGGAAAAGTATTTCCCCATG 2445  
 Qy 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420  
 Db 2446 CACGGAAATCTCATCTTTGGAAAAACAAGGCACAGGAACCTACCAATGTGGACATTGAATCA 2505

Qy 421 ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440  
 Db 2506 GTGCTTATTACAGACGAAGAATAATCAGAACGACTAATCCTGTGGCTACAGAACATAC 2565  
 Qy 441 GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis 460  
 Db 2566 GGACAGGTTGCCAACACCGTCAGAGTCAGAACACCACAGCTTCCTATGGAACTGTGGAC 2625  
 Qy 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 480  
 Db 2626 AGCCAGGGAATCTTACTCTGGAATGGTGTGGCAGGACCGCATGTCTATCTTCAAGGTCCC 2685  
 Qy 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500  
 Db 2686 ATTTGGGCCAAAACCTCTCACACGGACGGACACTTTCATCTCTCCGCTCATGGGAGGC 2745  
 Qy 501 PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520  
 Db 2746 TTTGGACTGAAACACCCCTCTCCCCAGATCCTGATCAAAAACACACCTGTGCCAGCGAAT 2805  
 Qy 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540  
 Db 2806 CCGCGGACCACTTTCACTCTCTGGAAAGTTTGTCTTCATCCAGTAGTATCCACCGGA 2865  
 Qy 541 GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro 560  
 Db 2866 CAGGTCAAGCTGGAAATAGAGTGGGAGCTGCAGAAAGAAAACAGCAACCGTGGAAACCCG 2925  
 Qy 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580  
 Db 2926 GAAATTCAGTACACCTCCAACACTACAACAAGTCGGTGAATGTGGAGTTTACCGTGGACGCA 2985  
 Qy 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599  
 Db 2986 AACGGTGTATTATTCTGAACCCCGCCCTATTGGCACTCGTTACCTTACCCGGAACCTTG 3042

RESULT 15

US-10-291-583-57  
 ; Sequence 57, Application US/10291583  
 ; Publication No. US20030138772A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gao, Guangping  
 ; APPLICANT: Wilson, James M.  
 ; APPLICANT: Alvira, Mauricio  
 ; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus  
 ; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1el Sequences Identif  
 ; FILE REFERENCE: UPN-02735USA  
 ; CURRENT APPLICATION NUMBER: US/10/291,583  
 ; CURRENT FILING DATE: 2002-11-12  
 ; PRIOR APPLICATION NUMBER: US 60/350,607  
 ; PRIOR FILING DATE: 2001-11-13  
 ; PRIOR APPLICATION NUMBER: US 60/341,117  
 ; PRIOR FILING DATE: 2001-12-17  
 ; PRIOR APPLICATION NUMBER: US 60/377,066  
 ; PRIOR FILING DATE: 2002-05-01  
 ; PRIOR APPLICATION NUMBER: US 60/386,675  
 ; PRIOR FILING DATE: 2002-06-05  
 ; NUMBER OF SEQ ID NOS: 120  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 57  
 ; LENGTH: 3123  
 ; TYPE: DNA  
 ; ORGANISM: new AAV serotype, clone A3.3  
 US-10-291-583-57

Alignment Scores:  
 Pred. No.: 1,71e-288 Length: 3123  
 Score: 2735.50 Matches: 493  
 Percent Similarity: 89.48% Conservative: 43  
 Best Local Similarity: 82.30% Mismatches: 62  
 Query Match: 84.14% Indels: 1  
 DB: 13 Gaps: 1





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 14, 2004, 16:55:45 ; Search time 496.857 Seconds  
(without alignments)  
3959.005 Million cell updates/sec

Title: US-09-807-802A-17  
Perfect score: 2906  
Sequence: 1 MASGGGAPMADNNEGADGVG.....NNGLYTEPRPIGTRYLTRPL 534

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2449703 seqs, 1841816367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09807802/runat\_11022004\_175610\_15987/app\_query.fasta\_1.2389  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09807802 @CGN 1 1 1100 @runat\_11022004\_175610\_15987  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq2:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result | Query | Score | Match | Length | DB | ID | Description |
|--------|-------|-------|-------|--------|----|----|-------------|
| -----  |       |       |       |        |    |    |             |

|    |        |       |      |    |                   |                   |
|----|--------|-------|------|----|-------------------|-------------------|
| 1  | 2906   | 100.0 | 4718 | 13 | US-10-291-583-6   | Sequence 6, Appli |
| 2  | 2906   | 100.0 | 7447 | 13 | US-10-216-870-11  | Sequence 11, Appl |
| 3  | 2536.5 | 87.3  | 3075 | 13 | US-10-291-583-26  | Sequence 26, Appl |
| 4  | 2528   | 87.0  | 4726 | 13 | US-10-291-583-8   | Sequence 8, Appli |
| 5  | 2520.5 | 86.7  | 3142 | 13 | US-10-291-583-25  | Sequence 25, Appl |
| 6  | 2518   | 86.6  | 8179 | 15 | US-10-205-942-5   | Sequence 5, Appli |
| 7  | 2481.5 | 85.4  | 4679 | 9  | US-09-804-898-1   | Sequence 1, Appli |
| 8  | 2481.5 | 85.4  | 4679 | 9  | US-09-945-681-10  | Sequence 10, Appl |
| 9  | 2481.5 | 85.4  | 4679 | 13 | US-10-136-819-6   | Sequence 6, Appli |
| 10 | 2481.5 | 85.4  | 4679 | 14 | US-10-038-972A-12 | Sequence 12, Appl |
| 11 | 2457.5 | 84.6  | 2205 | 13 | US-10-291-583-120 | Sequence 120, App |
| 12 | 2451.5 | 84.4  | 3123 | 13 | US-10-291-583-54  | Sequence 54, Appl |
| 13 | 2449.5 | 84.3  | 3122 | 13 | US-10-291-583-56  | Sequence 56, Appl |
| 14 | 2445.5 | 84.2  | 3113 | 13 | US-10-291-583-55  | Sequence 55, Appl |
| 15 | 2441   | 84.0  | 2489 | 13 | US-10-291-583-35  | Sequence 35, Appl |
| 16 | 2439.5 | 83.9  | 4680 | 13 | US-10-375-777-1   | Sequence 1, Appli |
| 17 | 2439.5 | 83.9  | 4680 | 14 | US-10-077-294-1   | Sequence 1, Appli |
| 18 | 2439.5 | 83.9  | 4680 | 14 | US-10-163-886-1   | Sequence 1, Appli |
| 19 | 2439.5 | 83.9  | 4680 | 15 | US-10-263-127-1   | Sequence 1, Appli |
| 20 | 2439   | 83.9  | 3098 | 13 | US-10-291-583-9   | Sequence 9, Appli |
| 21 | 2438.5 | 83.9  | 3123 | 13 | US-10-291-583-57  | Sequence 57, Appl |
| 22 | 2437.5 | 83.9  | 3106 | 13 | US-10-291-583-34  | Sequence 34, Appl |
| 23 | 2436   | 83.8  | 2495 | 13 | US-10-291-583-36  | Sequence 36, Appl |
| 24 | 2436   | 83.8  | 3098 | 13 | US-10-291-583-37  | Sequence 37, Appl |
| 25 | 2436   | 83.8  | 3106 | 13 | US-10-291-583-18  | Sequence 18, Appl |
| 26 | 2434.5 | 83.8  | 3123 | 13 | US-10-291-583-41  | Sequence 41, Appl |
| 27 | 2431.5 | 83.7  | 2504 | 13 | US-10-291-583-33  | Sequence 33, Appl |
| 28 | 2430   | 83.6  | 4721 | 13 | US-10-291-583-1   | Sequence 1, Appli |
| 29 | 2429.5 | 83.6  | 3129 | 13 | US-10-291-583-59  | Sequence 59, Appl |
| 30 | 2428.5 | 83.6  | 3128 | 13 | US-10-291-583-27  | Sequence 27, Appl |
| 31 | 2428.5 | 83.6  | 3197 | 13 | US-10-291-583-29  | Sequence 29, Appl |
| 32 | 2427.5 | 83.5  | 3084 | 13 | US-10-291-583-39  | Sequence 39, Appl |
| 33 | 2423.5 | 83.4  | 4675 | 10 | US-09-782-378A-1  | Sequence 1, Appli |
| 34 | 2423.5 | 83.4  | 4675 | 10 | US-09-782-378A-2  | Sequence 2, Appli |
| 35 | 2423.5 | 83.4  | 4675 | 13 | US-10-291-583-7   | Sequence 7, Appli |
| 36 | 2423.5 | 83.4  | 4675 | 15 | US-10-240-198-1   | Sequence 1, Appli |
| 37 | 2423   | 83.4  | 3098 | 13 | US-10-291-583-10  | Sequence 10, Appl |
| 38 | 2422.5 | 83.4  | 3113 | 13 | US-10-291-583-31  | Sequence 31, Appl |
| 39 | 2422.5 | 83.4  | 3128 | 13 | US-10-291-583-47  | Sequence 47, Appl |
| 40 | 2422   | 83.3  | 3106 | 13 | US-10-291-583-16  | Sequence 16, Appl |
| 41 | 2421.5 | 83.3  | 2501 | 13 | US-10-291-583-30  | Sequence 30, Appl |
| 42 | 2421.5 | 83.3  | 3121 | 13 | US-10-291-583-11  | Sequence 11, Appl |
| 43 | 2419.5 | 83.3  | 3095 | 13 | US-10-291-583-24  | Sequence 24, Appl |
| 44 | 2418.5 | 83.2  | 3095 | 13 | US-10-291-583-23  | Sequence 23, Appl |
| 45 | 2418   | 83.2  | 3102 | 13 | US-10-291-583-17  | Sequence 17, Appl |

ALIGNMENTS

RESULT 1  
US-10-291-583-6  
; Sequence 6, Application US/10291583  
; Publication No. US20030138772A1  
; GENERAL INFORMATION:  
; APPLICANT: Gao, Guangping  
; APPLICANT: Wilson, James M.  
; APPLICANT: Alvira, Mauricio  
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus  
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1el Sequences Identif  
; FILE REFERENCE: UPN-02735USA  
; CURRENT APPLICATION NUMBER: US/10/291,583  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: US 60/350,607  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/341,117  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: US 60/377,066  
; PRIOR FILING DATE: 2002-05-01  
; PRIOR APPLICATION NUMBER: US 60/386,675  
; PRIOR FILING DATE: 2002-06-05  
; NUMBER OF SEQ ID NOS: 120  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6

|  |      |  |               |  |      |  |  |  |  |
|--|------|--|---------------|--|------|--|--|--|--|
| ; LENGTH: 4718                                       |      |  |               |  |      |  |  |  |  |
| ; TYPE: DNA  |      |  |               |  |      |  |  |  |  |
| ; ORGANISM: adeno-associated virus serotype 1        |      |  |               |  |      |  |  |  |  |
| US-10-291-583-6                                      |      |  |               |  |      |  |  |  |  |
| Alignment Scores:                                    |      |  |               |  |      |  |  |  |  |
| Pred. NO.:   |      | 0  | Length:       |  | 4718 |  |  |  |  |
| Score:   |      | 2906.00  | Matches:      |  | 534  |  |  |  |  |
| Percent Similarity:                                  |      | 100.00%  | Conservative: |  | 0    |  |  |  |  |
| Best Local Similarity:                               |      | 100.00%  | Mismatches:   |  | 0    |  |  |  |  |
| Query Match:   |      | 100.00%  | Indels:       |  | 0    |  |  |  |  |
| DB:  |      | 13   | Gaps:         |  | 0    |  |  |  |  |
| US-09-807-802A-17 (1-534) x US-10-291-583-6 (1-4718) |      |  |               |  |      |  |  |  |  |
| QY   | 1    | MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly      | 20            |  |      |  |  |  |  |
| Db   | 2829 | ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGACGGAGTGGGT    | 2888          |  |      |  |  |  |  |
| QY   | 21   | AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr   | 40            |  |      |  |  |  |  |
| Db   | 2889 | AATGCCTCAGGAATTTGGCATTTGGATTCCACATGGCTGGCGACAGAGTCATCACCACC    | 2948          |  |      |  |  |  |  |
| QY   | 41   | SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer   | 60            |  |      |  |  |  |  |
| Db   | 2949 | AGCACCGCACCTGGGCCTTGCCACCTACAATAACCACTCTTACAAGCAAAATCTCCAGT    | 3008          |  |      |  |  |  |  |
| QY   | 61   | AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr   | 80            |  |      |  |  |  |  |
| Db   | 3009 | GCTTCAACGGGGCCGACGACGACAACCACTACTTCGGCTACAGCACCCCTGGGGGTAT     | 3068          |  |      |  |  |  |  |
| QY   | 81   | PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn   | 100           |  |      |  |  |  |  |
| Db   | 3069 | TTTGATTTCAACAGATTCCCACTGCCACTTTTCCACCGTACTGGCAGCGACTCATCAAC    | 3128          |  |      |  |  |  |  |
| QY   | 101  | AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys   | 120           |  |      |  |  |  |  |
| Db   | 3129 | AACAATTGGGGATTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAGTCAAG   | 3188          |  |      |  |  |  |  |
| QY   | 121  | GluValThrThrAsnAspGlyValThrIleAlaAsnAsnLeuThrSerThrValGln      | 140           |  |      |  |  |  |  |
| Db   | 3189 | GAGGTCAACGACGAATGATGGCGTCACAACCATCGCTAATAACCTTACCAGCACGGTTCAA  | 3248          |  |      |  |  |  |  |
| QY   | 141  | ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys   | 160           |  |      |  |  |  |  |
| Db   | 3249 | GTCTTCTCGGACTCGGAGTACAGCTTCCGTACGTCTCTCGGCTCTGCGCACCAAGGCTGC   | 3308          |  |      |  |  |  |  |
| QY   | 161  | LeuProPropheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn   | 180           |  |      |  |  |  |  |
| Db   | 3309 | CTCCCTCCGTTCCCGCGGACGTGTTTCATGATTCGCAATACGGCTACCTGACGCTCAAC    | 3368          |  |      |  |  |  |  |
| QY   | 181  | AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln   | 200           |  |      |  |  |  |  |
| Db   | 3369 | AATGGCAGCCCAAGCCGTGGACGTTTCATCTCTTTTACTGCCTGGAATATTTCCCTTCTCAG | 3428          |  |      |  |  |  |  |
| QY   | 201  | MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPropheHis   | 220           |  |      |  |  |  |  |
| Db   | 3429 | ATGCTGAGAACGGGCAACAACTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC   | 3488          |  |      |  |  |  |  |
| QY   | 221  | SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr   | 240           |  |      |  |  |  |  |
| Db   | 3489 | AGCAGCTACGCGCACAGCCAGAGCCTGGACCGGCTGATGAATCTCTCATCGACCAATAC    | 3548          |  |      |  |  |  |  |
| QY   | 241  | LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu   | 260           |  |      |  |  |  |  |
| Db   | 3549 | CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAAGGACTTGCTG   | 3608          |  |      |  |  |  |  |
| QY   | 261  | PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro   | 280           |  |      |  |  |  |  |
| Db   | 3609 | TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTACGCCCAAAACTGGCTACCTGGACCC    | 3668          |  |      |  |  |  |  |
| QY   | 281  | CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnSerAsnPheThr      | 300           |  |      |  |  |  |  |

|  |      |  |      |
|--|------|--|------|
| Db   | 3669 | TGTTATCGGCAGCAGCGCGTCTTCTAAAAACAAAAACAGACAACAACAGCAATTTTACC    | 3726 |
| QY   | 301  | TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr   | 320  |
| Db   | 3729 | TGGACTGGTGTCTTCAAAATATATAACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT | 3788 |
| QY   | 321  | AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle   | 340  |
| Db   | 3789 | GCTATGGCCTCACACAAAGACGACGAGACAAAGTCTTTTCCCATGAGCGGTGCATGATT    | 3848 |
| QY   | 341  | PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp   | 360  |
| Db   | 3849 | TTTGGAAAAGAGAGCGCGGAGCTTCAAACACTGTCATTTGGACAATGTATGATTACAGAC   | 3908 |
| QY   | 361  | GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal   | 380  |
| Db   | 3909 | GAAGAGGAAATTAAGCCACTAACCCCTGTGGCCACCGAAAGATTTGGGACCGTGCAGTC    | 3968 |
| QY   | 381  | AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu   | 400  |
| Db   | 3969 | AATTTCCAGAGCAGCAGCAGACCCCTGCGACCGGAGATGTGCATGTCTATGGGAGCATTA   | 4028 |
| QY   | 401  | ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle   | 420  |
| Db   | 4029 | CCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGTCAGGGTCCCATTGCGGCCAAAATT  | 4088 |
| QY   | 421  | ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn   | 440  |
| Db   | 4089 | CCTCACACAGATGGACACTTTCACCCGCTCTCTCTTATGGCGGCTTTGGACTCAAGAAC    | 4148 |
| QY   | 441  | ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe   | 460  |
| Db   | 4149 | CGCCTCTCTCAGATCTCTCATCAAAAACACGCGCTGTCTCTCGAATCCTCCGCGGAGTTT   | 4208 |
| QY   | 461  | SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu   | 480  |
| Db   | 4209 | TCAGCTACAAAGTTTGTCTCATTCATCACCCCATACTCCACAGGACAAGTGAGTGTGGAA   | 4268 |
| QY   | 481  | IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr   | 500  |
| Db   | 4269 | ATTGAATGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAATCCCAGAGTGCAGTACACA    | 4328 |
| QY   | 501  | SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr   | 520  |
| Db   | 4329 | TCCAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAACAATGGACTTTTACT    | 4388 |
| QY   | 521  | GluProArgProIleGlyThrArgTyrLeuThrArgProLeu                     | 534  |
| Db   | 4389 | GAGCCTCGCCCATTTGGCACCCCGTTACCTTACCCGTCCTG 4430                 |      |
| RESULT 2   |      |  |      |
| US-10-216-870-11   |      |  |      |
| ; Sequence 11, Application US/10216870                                     |      |  |      |
| ; Publication No. US20030148506A1  |      |  |      |
| ; GENERAL INFORMATION:   |      |  |      |
| ; APPLICANT: KOTIN, ROBERT M   |      |  |      |
| ; APPLICANT: URABE, MASASHI  |      |  |      |
| ; APPLICANT: DING, CHUAN-TIAN  |      |  |      |
| ; TITLE OF INVENTION: PRODUCTION OF ADENO-ASSOCIATED VIRUS IN INSECT CELLS |      |  |      |
| ; FILE REFERENCE: 402133   |      |  |      |
| ; CURRENT APPLICATION NUMBER: US/10/216,870                                |      |  |      |
| ; CURRENT FILING DATE: 2002-08-13  |      |  |      |
| ; NUMBER OF SEQ ID NOS: 24   |      |  |      |
| ; SOFTWARE: PatentIn version 3.1   |      |  |      |
| ; SEQ ID NO 11   |      |  |      |
| ; LENGTH: 7447   |      |  |      |
| ; TYPE: DNA  |      |  |      |
| ; ORGANISM: Artificial Sequence  |      |  |      |
| ; FEATURE:   |      |  |      |
| ; OTHER INFORMATION: Synthetic   |      |  |      |
| US-10-216-870-11   |      |  |      |
| Alignment Scores:  |      |  |      |

RESULT 2  
US-10-216-870-11  
; Sequence 11, Application US/10216870  
; Publication No. US20030148506A1  
; GENERAL INFORMATION:  
; APPLICANT: KOTIN, ROBERT M  
; APPLICANT: URABE, MASASHI  
; APPLICANT: DING, CHUAN-TIAN  
; TITLE OF INVENTION: PRODUCTION OF ADENO-ASSOCIATED VIRUS IN INSECT CELLS  
; FILE REFERENCE: 402133  
; CURRENT APPLICATION NUMBER: US/10/216,870  
; CURRENT FILING DATE: 2002-08-13  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 11  
; LENGTH: 7447  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
US-10-216-870-11  
Alignment Scores:

|   |         |  |      |
|---|---------|--|------|
| Pred. No.:  | 0       | Length:  | 7447 |
| Score:  | 2906.00 | Matches:   | 534  |
| Percent Similarity:                                   | 100.00% | Conservative:  | 0    |
| Best Local Similarity:                                | 100.00% | Mismatches:  | 0    |
| Query Match:  | 100.00% | Indels:  | 0    |
| DB:   | 13      | Gaps:  | 0    |
| US-09-807-802A-17 (1-534) x US-10-216-870-11 (1-7447) |         |  |      |
| QY  | 1       | MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly      | 20   |
| Db  | 5227    | ATGGCTTCAGGCGGTGGCGCACCACCAATGGCAGACAATAACGAAGCGCCGACGGAGTGGGT | 5286 |
| QY  | 21      | AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr   | 40   |
| Db  | 5287    | AATGCCTCAGGAAATTGGCATTGGCATTCCACATGGCTGGCGCAGAGTCATCACACC      | 5346 |
| QY  | 41      | SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer   | 60   |
| Db  | 5347    | AGCACCCGCACCTGGGCCTTGCCACCTACAAATAACCACTCTACAAGCAAAATCTCCAGT   | 5406 |
| QY  | 61      | AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr   | 80   |
| Db  | 5407    | GCTTCAACGGGGCCAGCAACGACAACCACTACTTCGCTACAGCACCCCTCGGGGTAT      | 5466 |
| QY  | 81      | PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn   | 100  |
| Db  | 5467    | TTTGATTCAACAGATTCCACTGCCACTTTTCACCACGTGACTGGCAGCGACTCATCAAC    | 5526 |
| QY  | 101     | AsnAsnTrpGlyPheArgProLysArgLeuAsnPhelLysLeuPheAsnIleGlnValLys  | 120  |
| Db  | 5527    | AACAATTGGGGATTCCGGCCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAGTCAAG   | 5586 |
| QY  | 121     | GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln   | 140  |
| Db  | 5587    | GAGGTACAGACGATGATGGGCTCAACACCATCGCTAATAAACCTTACCAGCACGGTTCAA   | 5646 |
| QY  | 141     | ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys   | 160  |
| Db  | 5647    | GTCTTCTCGGACTCGGAGTACAGACTTCGTAACGCTCCGGCTCTGCGCACCGGGCTGC     | 5706 |
| QY  | 161     | LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn   | 180  |
| Db  | 5707    | CTCCCTCCGTTCCCGCGGACGTGTTTCATGATTCCGCAATACGGCTACCTGACGTTCAAC   | 5766 |
| QY  | 181     | AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln   | 200  |
| Db  | 5767    | AATGGCAGCCAAAGCCGTGGGACGTTTCATCTTTTACTGCTGGAATATTTCCCTTCTCAG   | 5826 |
| QY  | 201     | MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis   | 220  |
| Db  | 5827    | ATGCTGAGAACGGGCAACAATTACCTTCAGCTACACTTTGAGGAAGTGCTTTCCAC       | 5886 |
| QY  | 221     | SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr   | 240  |
| Db  | 5887    | AGCAGCTACGCGCACAGCCAGAGCCTGGACCGGTGATGAATCCTCTCATCGACCAATAC    | 5946 |
| QY  | 241     | LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu   | 260  |
| Db  | 5947    | CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAAGGACTTGCTG   | 6006 |
| QY  | 261     | PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro   | 280  |
| Db  | 6007    | TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAACTGGCTACCTGGACCC   | 6066 |
| QY  | 281     | CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr   | 300  |
| Db  | 6067    | TGTTATCGGCAGCAGCGCGTTTCTAAACAACAAACAGACAACAACAGCAATTTTACC      | 6126 |
| QY  | 301     | TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr   | 320  |
| Db  | 6127    | TGGACTGGTGTCTCAAAATATAACTCAATGGGCGTGAATCCATCATCAACCCCTGGCACT   | 6186 |

|    |      |  |      |
|----|------|--|------|
| QY | 321  | AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle   | 340  |
| Db | 6187 | GCTATGGCCTCACACAAAGACGACGAAGACAAGTTCTTTCCCATGAGCGGTGTCATGATT   | 6246 |
| QY | 341  | PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp   | 360  |
| Db | 6247 | TTTGAAAAGAGAGCGCCGGAGCTTCAAAACACTGCATTTGGACAATGTTCATGATTACAGAC | 6306 |
| QY | 361  | GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal   | 380  |
| Db | 6307 | GAAGAGGAAATTAAAGCCACTAACCCCTGTGGCCACCAGAAAGATTTTGGACCGTGGCAGTC | 6366 |
| QY | 381  | AsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu      | 400  |
| Db | 6367 | AATTTCCAGAGCAGCAGCACAGACCCTGCGACCGGAGATGTGCATGCTATGGAGCATTA    | 6426 |
| QY | 401  | ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle   | 420  |
| Db | 6427 | CCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCAGGTCCCATTTTGGGCCAAAATT   | 6486 |
| QY | 421  | ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn   | 440  |
| Db | 6487 | CCTCACACAGATGGACACTTTTCACCCGCTCTCTCTTATGGGGGGCTTTGGACTCAAGAAC  | 6546 |
| QY | 441  | ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe   | 460  |
| Db | 6547 | CCGCCTCCTCAGATCCTCATCAAAAACACGCTGTTCTCGGAATCCTCCGCGGAGTTT      | 6606 |
| QY | 461  | SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu   | 480  |
| Db | 6607 | TCAGCTACAAAGTTGCTTTCATTTCATCACCCAATACTCCACAGGACAAGTGAAGTGGAA   | 6666 |
| QY | 481  | IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr   | 500  |
| Db | 6667 | ATTGAATGGGAGCTGCAGAAAGAAAACAGCAAGCGTGAATCCCGAAGTGCAGTACACA     | 6726 |
| QY | 501  | SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr   | 520  |
| Db | 6727 | TCCAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAACAATGGACTTTTATACT  | 6786 |
| QY | 521  | GluProArgProIleGlyThrArgTyrLeuThrArgProLeu                     | 534  |
| Db | 6787 | GAGCCTCGCCCCCATTTGGCACCCGTTTACCTTACCTTACCCGTTCCCTG             | 6828 |

RESULT 3

US-10-291-583-26  
; Sequence 26, Application US/10291583  
; Publication No. US20030138772A1  
; GENERAL INFORMATION:  
; APPLICANT: Gao, Guangping  
; APPLICANT: Wilson, James M.  
; APPLICANT: Alvira, Mauricio  
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus  
; FILE REFERENCE: UPN-02735USA  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR FILING DATE: 2002-05-01  
; PRIOR FILING DATE: 2002-06-05  
; NUMBER OF SEQ ID NOS: 120  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 26  
; LENGTH: 3075  
; TYPE: DNA  
; ORGANISM: new AAV serotype, clone H2  
US-10-291-583-26



Alignment Scores:

Pred. No.: 1.49e-285 Length: 3075  
Score: 2536.50 Matches: 455  
Percent Similarity: 91.57% Conservative: 34  
Best Local Similarity: 85.21% Mismatches: 44  
Query Match: 87.28% Indels: 1  
DB: 13 Gaps: 1

US-09-807-802A-17 (1-534) x US-10-291-583-26 (1-3075)

|    |      |  |      |
|----|------|--|------|
| Qy | 1    | MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly        | 20   |
| Db | 1409 | ATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGGTCCGATGGAGTGGGT      | 1468 |
| Qy | 21   | AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThr        | 40   |
| Db | 1469 | AAATCCTCAGGAAATTTGGCATTTGCCATTTCCCAATGGCTGGGCGACAGAGTCATCACCACC  | 1528 |
| Qy | 41   | SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSer        | 60   |
| Db | 1529 | AGCACCCGAACCTGGGCCCTGCCACATACAACAACCACTTACAAGCAAATCTCCAGC        | 1588 |
| Qy | 61   | AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTipGlyTyr     | 80   |
| Db | 1589 | CAATCA---GGAGCCAGCAACGACACCACTACTTTGGCTACAGCACCCCTGGGGGTAT       | 1645 |
| Qy | 81   | PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn     | 100  |
| Db | 1646 | TTTGACTTCAACAGATTCCCACTGCCACTTTTCACCACGTGATGGCAAGACTCATCAAC      | 1705 |
| Qy | 101  | AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys     | 120  |
| Db | 1706 | AACAAC TGGGGATTCCGGCCCCAAAAGACTCAACTTCAAGCTCTTTAATATTCAAGTCAA    | 1765 |
| Qy | 121  | GluValThrThrAsnAspGlyValThrIleAlaAsnAsnLeuThrSerThrValGln        | 140  |
| Db | 1766 | GAGGTACCGAGAATGACGGTACGACGACGATTGCCAATAACTTACCAGCACGGTTTCAG      | 1825 |
| Qy | 141  | ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys     | 160  |
| Db | 1826 | GTGTTTACTGACTCGGAGTACCAGCTCCGTACGTCCTGGCTCGGCGCATCAAGGATGC       | 1885 |
| Qy | 161  | LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn     | 180  |
| Db | 1886 | CTCCCGCCGTTTCCAGCGGACGCTCTTCATGGTCCACAGTATGGATACCTCACCCCTGAAC    | 1945 |
| Qy | 181  | AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln     | 200  |
| Db | 1946 | AACGGGAGTACGGGGTAGGACGCTCTTCTTTACTGCTGGAGTACTTTCTTCTCTCAG        | 2005 |
| Qy | 201  | MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis     | 220  |
| Db | 2006 | ATGTCGCGTACTTGGAAACAACACTTTTCAGTTTCAGCTACACTTTTGAAGACGTGCCTTCCAC | 2065 |
| Qy | 221  | SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr     | 240  |
| Db | 2066 | AGCAGCTACGCTCACAGCCAGAGTCTGGATCGGCTGATGAATCCTCTGATCGACCAGTAC     | 2125 |
| Qy | 241  | LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu     | 260  |
| Db | 2126 | CTGTATTATCTGAACAAGACACAAACAATAAGTGGAACTCTTCAGCAGCTCTCGGCTACTG    | 2185 |
| Qy | 261  | PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro     | 280  |
| Db | 2186 | TTTAGCCAAGCTGGACCAACCAACATGTCTCTTCAAGCTAAAACTGGCTGCCTGGACCT      | 2245 |
| Qy | 281  | CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr     | 300  |
| Db | 2246 | TGCTACAGACAGCAGCGTCTGTCAAAAACAGGCAACGACAAACAACACAGCAACTTTCCC     | 2305 |
| Qy | 301  | TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr     | 320  |
| Db | 2306 | TGGACTGCAGCTACAAAGATATCATCTAAATGGCCGGGACTCGTTGGTTAATCCAGGACCA    | 2365 |

|    |      |   |      |
|----|------|---|------|
| Qy | 321  | AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle    | 340  |
| Db | 2366 | GCTATGGCCAGTCACAAGGATGACGAAGAAAAGTTTTTCCCCATGCATGGAACCCCTGATA   | 2425 |
| Qy | 341  | PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp    | 360  |
| Db | 2426 | TTTGGTAAACAAGGAACAAATGCCAACGACCGGATTTGGAAAATGTCTATGATTACAGAT    | 2485 |
| Qy | 361  | GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal    | 380  |
| Db | 2486 | GAAGAAGAAATCAGGGCCACCATCCCGTGGCTACGGAGCAGTACGGGACTGTGTCAAAT     | 2545 |
| Qy | 381  | AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu    | 400  |
| Db | 2546 | AATTTGCAAAACTCAAAACACTGGTCCAAC TACTGGAAC TGTCAATCGCCAAGGCGTTA   | 2605 |
| Qy | 401  | ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle    | 420  |
| Db | 2606 | CCTGGTATGGTGTGCAGGATCGAGACGTGTACTGCAGGGACCCCATTTGGGCCAAGATT     | 2665 |
| Qy | 421  | ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn    | 440  |
| Db | 2666 | CCTCACACCGATGGACACTTTTCATCCTTCTCCACTGATGGAGGTTTTTGGACTCAAACAC   | 2725 |
| Qy | 441  | ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe    | 460  |
| Db | 2726 | CCGCCTCCTCAGATCATGATCAAAAACACTCCCGTCCAGCCAATCCTCCCAACAACCTTC    | 2785 |
| Qy | 461  | SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu    | 480  |
| Db | 2786 | AGTTCTGCCAAGTTTGGCTTCTTTTCATCACACAGTATTCACGGGACAGGTCAGCGTGGAG   | 2845 |
| Qy | 481  | IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr    | 500  |
| Db | 2846 | ATCGAGTGGGAGCTGCAGAAGGAGAACACAGCAAAACGCTGGAATCCCGAAATTCAGTACACT | 2905 |
| Qy | 501  | SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr    | 520  |
| Db | 2906 | TCCAAC TACAACAAGTCTGTTAATGTGGACTTTTACTGTGGACACTAATGGTGTGTATTCA  | 2965 |
| Qy | 521  | GluProArgProIleGlyThrArgTyrLeuThrArgProLeu                      | 534  |
| Db | 2966 | GAGCCTCGCCCCCATTTGGCACCAGATACCTGACTCGTAATCTG                    | 3007 |

RESULT 4

US-10-291-583-8  
; Sequence 8, Application US/10291583  
; Publication No. US20030138772A1  
; GENERAL INFORMATION:  
; APPLICANT: Gao, Guangping  
; APPLICANT: Wilson, James M.  
; APPLICANT: Alvira, Mauricio  
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus  
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1e1 Sequences Identif  
; FILE REFERENCE: UPN-02735USA  
; CURRENT APPLICATION NUMBER: US/10/291,583  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: US 60/350,607  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/341,117  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: US 60/377,066  
; PRIOR FILING DATE: 2002-05-01  
; PRIOR APPLICATION NUMBER: US 60/386,675  
; PRIOR FILING DATE: 2002-06-05  
; NUMBER OF SEQ ID NOS: 120  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 4726  
; TYPE: DNA  
; ORGANISM: adeno-associated virus serotype 3  
US-10-291-583-8

|  |           |   |      |  |
|--|-----------|---|------|--|
| Alignment Scores:                                    |           |   |      |  |
| Pred. No.:   | 2.93e-284 | Length:   | 4726 |  |
| Score:   | 2528.00   | Matches:  | 459  |  |
| Percent Similarity:                                  | 91.03%    | Conservative:   | 28   |  |
| Best Local Similarity:                               | 85.79%    | Mismatches:   | 46   |  |
| Query Match:   | 86.99%    | Indels:   | 2    |  |
| DB:  | 13        | Gaps:   | 2    |  |
| US-09-807-802A-17 (1-534) x US-10-291-583-8 (1-4726) |           |   |      |  |
| QY   | 1         | MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly         | 20   |  |
| DB   | 2815      | ATGGCTTCAGCGGTGGCGCAACCAATGGCAGACAAATAACGAGGGTGCCGATGGAGTGGT      | 2874 |  |
| QY   | 21        | AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr      | 40   |  |
| DB   | 2875      | AATTCCTCAGGAAATTGGCAFTGCGATTCCCAATGGCTGGCGCAGAGTCATCACACC         | 2934 |  |
| QY   | 41        | SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer      | 60   |  |
| DB   | 2935      | AGCACAGAACTGGGCCCTGCCACTTACAACAACCACTCTACAAGCAAAATCTCCAGC         | 2994 |  |
| QY   | 61        | AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr      | 80   |  |
| DB   | 2995      | CAATCA---GGAGCTTCAACGACAACCACTACTTTGGGTACAGCACCCCTTGGGGGTAT       | 3051 |  |
| QY   | 81        | PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn      | 100  |  |
| DB   | 3052      | TTTGACTTTAACAGATTCCACTGCCACTTCTCACCACTGACTGGCAGCGACTCATTAAC       | 3111 |  |
| QY   | 101       | AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys      | 120  |  |
| DB   | 3112      | AACAACATGGGGATTCCGGCCCAAGAAACTCAGCTTCAAGCTCTTCAACATCCAGTTAGA      | 3171 |  |
| QY   | 121       | GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln      | 140  |  |
| DB   | 3172      | GGGGTCACGCAGAACGATGGCAGCAGCACTATTGCCAATAACCTTACCAGCACGGTTCAA      | 3231 |  |
| QY   | 141       | ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys      | 160  |  |
| DB   | 3232      | GTGTTTACGGACTCGGAGTATCAGCTCCCGTACGTGCTCGGGTCGGCGCACCAAGGCTGT      | 3291 |  |
| QY   | 161       | LeuProPropheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn      | 180  |  |
| DB   | 3292      | CTCCCGCGGTTTCCAGCGGACGCTTCTCATGTTGCTCCCTCAGTATGGATACCTCACCCCTGAAC | 3351 |  |
| QY   | 181       | AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln      | 200  |  |
| DB   | 3352      | AACGGAAGTCAAGCGGTGGGACGCTCATCTTTTACTGCTCGGAGTACTTCCCTTCGCAG       | 3411 |  |
| QY   | 201       | MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis      | 220  |  |
| DB   | 3412      | ATGCTAAGGACTGGAAATAACTTCCAATTTCAGCTATACCTTCGAGGATGTACCTTTTCAC     | 3471 |  |
| QY   | 221       | SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr      | 240  |  |
| DB   | 3472      | AGCAGCTACGCTCACAGCCAGAGTTTGGATCGCTTGATGAATCTCTTATTGATCAGTAT       | 3531 |  |
| QY   | 241       | LeuTyrTyrLeuAsnArgThrGln--AsnGlnSerGlySerAlaGlnAsnLysAspLeu       | 259  |  |
| DB   | 3532      | CTGTACTACCTGAACAGAACGCAAGCAACCACTCTGGAACAACCAACCAATCACGSGTG       | 3591 |  |
| QY   | 260       | LeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGly      | 279  |  |
| DB   | 3592      | CTTTTATGCCAGGCTGGGCCCTCAGTCTATGTCTTTGCGAGGCCAGAAATTGGCTACCTGGG    | 3651 |  |
| QY   | 280       | ProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPhe      | 299  |  |
| DB   | 3652      | CCCTGCTACCGGCAACAGAGACTTTCAAAGACTGTCTAACGCAACAACAACAGTAACTTT      | 3711 |  |
| QY   | 300       | ThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGly      | 319  |  |

|   |      |  |      |  |
|---|------|--|------|--|
| DB  | 3712 | CCTTGGACAGCGGCCAGCAAAATATCATCTCAATGGCCGCGACTCGCTGGTGAATCCAGGA  | 3771 |  |
| QY  | 320  | ThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMet   | 339  |  |
| DB  | 3772 | CCAGCTATGGCCAGTCAACAGGACGATGAAGAAAAATTTTCCCTATGACGGCAACTA      | 3831 |  |
| QY  | 340  | IlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThr   | 359  |  |
| DB  | 3832 | ATATTTGGCAAAAGGAGGCAACGCAAGTAACGCAGAAATTAGATAATGTAAATGATTACG   | 3891 |  |
| QY  | 360  | AspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAla   | 379  |  |
| DB  | 3892 | GATGAAGAGAGATTCTGTAACCACCAATCTGTGGCAACAGAGCAGTATGGAACCTGTGGA   | 3951 |  |
| QY  | 380  | ValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAla   | 399  |  |
| DB  | 3952 | AATAACTTGCAGAGCTCAAATACAGCTCCACGACTGGAACCTGTCAATCATCATCAGGGGCC | 4011 |  |
| QY  | 400  | LeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLys   | 419  |  |
| DB  | 4012 | TTACCTGGCATGGTGTGGCAAGATCGTGACGTGTACTCTCAAGGACCTATCTGGGCAAG    | 4071 |  |
| QY  | 420  | IleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLys   | 439  |  |
| DB  | 4072 | ATTCTCACAGGATGGACACTTTCATCTCTCTCTGATGGGAGGCTTTGGACTGAAA        | 4131 |  |
| QY  | 440  | AsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGlu      | 459  |  |
| DB  | 4132 | CATCCGCTCTCTCAAATCATGATCAAAAAATACTCCGGTACCGGCAAAATCCTCCGACGACT | 4191 |  |
| QY  | 460  | PheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerVal   | 479  |  |
| DB  | 4192 | TTCAGCCCCGCCAAGTTTGCTTCATTTATCACTCAGTACTCCACTGGACAGGTTCAGCGTG  | 4251 |  |
| QY  | 480  | GluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTyrAsnProGluValGlnTyr   | 499  |  |
| DB  | 4252 | GAAATTGAGTGGGAGCTACAGAAAGAAACAGCAAAACGTTGGAATCCAGAGATTTCAGTAC  | 4311 |  |
| QY  | 500  | ThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyr   | 519  |  |
| DB  | 4312 | ACTTCCAACCTACAACAAGTCTGTTAATGTGGACTTTACTGTAGACACTAATGGTGTATTAT | 4371 |  |
| QY  | 520  | ThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu                  | 534  |  |
| DB  | 4372 | AGTGAACCTCGCCCTATTGGAACCCCGGTATCTCACACGAAACTTG                 | 4416 |  |
| RESULT 5  |      |  |      |  |
| US-10-291-583-25/c  |      |  |      |  |
| ; Sequence 25, Application US/10291583  |      |  |      |  |
| ; Publication No. US20030138772A1   |      |  |      |  |
| ; GENERAL INFORMATION:  |      |  |      |  |
| ; APPLICANT: Gao, Guangping   |      |  |      |  |
| ; APPLICANT: Wilson, James M.   |      |  |      |  |
| ; APPLICANT: Alvira, Mauricio   |      |  |      |  |
| ; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus |      |  |      |  |
| ; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1                     |      |  |      |  |
| ; FILE REFERENCE: UPN-02735USA  |      |  |      |  |
| ; CURRENT APPLICATION NUMBER: US/10/291,583   |      |  |      |  |
| ; CURRENT FILING DATE: 2002-11-12   |      |  |      |  |
| ; PRIOR APPLICATION NUMBER: US 60/350,607   |      |  |      |  |
| ; PRIOR FILING DATE: 2001-11-13   |      |  |      |  |
| ; PRIOR APPLICATION NUMBER: US 60/341,117   |      |  |      |  |
| ; PRIOR FILING DATE: 2001-12-17   |      |  |      |  |
| ; PRIOR APPLICATION NUMBER: US 60/377,066   |      |  |      |  |
| ; PRIOR FILING DATE: 2002-05-01   |      |  |      |  |
| ; PRIOR APPLICATION NUMBER: US 60/386,675   |      |  |      |  |
| ; PRIOR FILING DATE: 2002-06-05   |      |  |      |  |
| ; NUMBER OF SEQ ID NOS: 120   |      |  |      |  |
| ; SOFTWARE: PatentIn version 3.1  |      |  |      |  |
| ; SEQ ID NO 25  |      |  |      |  |
| ; LENGTH: 3142  |      |  |      |  |
| ; TYPE: DNA   |      |  |      |  |
| ; ORGANISM: new AAV serotype, clone H6  |      |  |      |  |

US-10-291-583-25

Alignment Scores:

Pred. No.: 1.15e-283 Length: 3142  
Score: 2520.50 Matches: 453  
Percent Similarity: 91.39% Conservative: 35  
Best Local Similarity: 84.83% Mismatches: 45  
Query Match: 86.73% Indels: 1  
DB: 13 Gaps: 1

US-09-807-802A-17 (1-534) x US-10-291-583-25 (1-3142)

QY 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20  
Db 1733 ATGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGGTCCGATGGAGTGGT 1674  
QY 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40  
Db 1673 AATTCTCAGGAATTGGCATTTGGATTCCCAATGGCTGGCGCAGAGTCATCACCACC 1614  
QY 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
Db 1613 AGCACCAGAACCTGGGCCCTGCCACATACAACAACCACTCTACAAGCAATCTCCAGC 1554  
QY 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80  
Db 1553 CAATCA--GGAGCCAGCAACGACAACCACTACTTTGGCTACAGCACCCCTGGGGGTAT 1497  
QY 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100  
Db 1496 TTTGACTTCAACAGATTCCACTGCCACTTTTACCACGCTGACTGGCAAGACTCGTCAAC 1437  
QY 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120  
Db 1436 AACAAACCGGGATTCCGGCCCCAAAGACTCAACTTCAAGCTCTTAATATTCAAGTCAAA 1377  
QY 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140  
Db 1376 GAGGTACGCAGAAATGACGGTACGACGACGATTGCCAATAACCTTACCAGCACGGTTT 1317  
QY 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160  
Db 1316 GTGTTTACTGACTCGGAGTACCAGCTCCGTACGTCTGGGGTCTGGCGCATCAAGGATGC 1257  
QY 161 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
Db 1256 CTCCCGCCGTTTCCAGCGGACGCTCTTCATGGTCCCACAGTATGGATACCTCACCCCTGAAC 1197  
QY 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
Db 1196 AACGGGAGTCAGCGGGTAGGACGCTCTTCTTTTACTGCTGGAGTACTTTCTTCTCTCAG 1137  
QY 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPropheHis 220  
Db 1136 ATGCTGCGTACTGGAACAACAATTTTCAGTTTCCAGTACAGTATTTGAAGACGTGCCTTTCCAC 1077  
QY 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
Db 1076 AGCAGCTACGCTACAGCCAGAGTCTGGATCGGCTGATGAATCTCTGATCGACCATAC 1017  
QY 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260  
Db 1016 CTGTATTATCTGAACAAGACACAACAATAAGTAGGAACTCTTCAGCAGTCTCGGCTACTG 957  
QY 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280  
Db 956 TTTAGCCAAGCTGGACCAACCAACATGTCTCTCAAGCTAAACCTGGCTGCCTGGACCT 897  
QY 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
Db 896 TGCTACAGACAGCAGCGTCTGTCAAAACAGGCAACGACAAACAACAGCAACTTTTCCC 837  
QY 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320

Db 836 TGGACTGCAGCTACAAAGTATCATCTAAATGGCCGGGACTCGTTGGTTAATCCAGGACCA 777  
QY 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340  
Db 776 GCTATGGCCAGTCACAAGGATGACGAAGAAAAGTTTTTCCCCATGTCATGGAACCCCTGATA 717  
QY 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360  
Db 716 TTTGGTAAACAAAGGAACAATGCCAACGACGCGGATTTGGAAAATGTCAATGATTACAGAT 657  
QY 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380  
Db 656 GAAGAAGAAATCAGGCCCACTCAATCCCGTGGCTACGGAGCAGTACGGGACTGTGTCAAAT 597  
QY 381 AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400  
Db 596 AATTGCAAAACTCAACACTGGTCCAACTACTGGAACGTGCAATCACCACAGGAGCGTTA 537  
QY 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420  
Db 536 CCTGGTATGGTGTGGCAGGATCGAGACGTTGACCTGCAGGGACCCATTTGGGCCAAGATT 477  
QY 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440  
Db 476 CCTCACACCGATGGACACTTTTCATCCTTCTCCTGATGGAGGTTTTGGACTCAACACAC 417  
QY 441 ProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460  
Db 416 CCGCTCTCTCAGATCATGATCAAAAACACTCCGTTCCAGCCAATCTCCCAAAACTTC 357  
QY 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480  
Db 356 AGTTCTGCCAAGTTTGTCTTTCATCACACAGTATTCACGGGACAGGTCAGCGTGGAG 297  
QY 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 500  
Db 296 ATCGAGTGGGAGCTGCAGAAGGAGAACAGCAACGCTGGAATCCCGAAATTCAGTACACT 237  
QY 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520  
Db 236 TCCAACACTACAACAGTCTGTTAATGTGGACTTTACTGTGGACACATAATGGTGTGTATTCA 177  
QY 521 GluProArgProIleGlyThrArgTyrLeuThrLeuThrArgProLeu 534  
Db 176 GAGCCTCGCCCCATTGGCACCAGATACCTGACTCGTAAATCTG 135

RESULT 6

US-10-205-942-5  
; Sequence 5, Application US/10205942  
; Publication No. US20030053990A1  
; GENERAL INFORMATION:  
; APPLICANT: University of No. US20030053990A1th Carolina-Chapel Hill  
; APPLICANT: Rabinowitz, Joseph E.  
; APPLICANT: Samulski, Richard J.  
; APPLICANT: Xiao, Weidong  
; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING THE SAME  
; FILE REFERENCE: 5470-186  
; CURRENT APPLICATION NUMBER: US/10/205,942  
; CURRENT FILING DATE: 2002-07-26  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 8179  
; TYPE: DNA  
; ORGANISM: Adeno-associated virus  
US-10-205-942-5

Alignment Scores:  
Pred. No.: 1.04e-282 Length: 8179  
Score: 2518.00 Matches: 458  
Percent Similarity: 90.65% Conservative: 27  
Best Local Similarity: 85.61% Mismatches: 48



|  |        |   |      |
|--|--------|---|------|
| Query Match:   | 86.65% | Indels:   | 2    |
| DB:  | 15     | Gaps:   | 2    |
| US-09-807-802A-17 (1-534) x US-10-205-942-5 (1-8179) |        |   |      |
| QY   | 1      | MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly     | 20   |
| Db   | 2739   | ATGGCTACAGGCAGTGGCGCAACCAATGGCAGACAATAACGAGGGCGCGCGAGTGGGT    | 2798 |
| QY   | 21     | AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr  | 40   |
| Db   | 2799   | AATTCTCCGGAATTTGGCATTTGGATTCCCATGGATGGGCGACAGATCATCAACCACC    | 2858 |
| QY   | 41     | SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer  | 60   |
| Db   | 2859   | AGCACCGGAACCTGGGCCCTGGCCACCTACAACAACCACTCTACAAACAAATTTCCAGC   | 2918 |
| QY   | 61     | AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr  | 80   |
| Db   | 2919   | CAATCA---GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTGGGGGTAT  | 2975 |
| QY   | 81     | PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn  | 100  |
| Db   | 2976   | TTTGACTTCAACAGATTCCACTGCCACTTTTCACCACCGTACTGGCAAGACTCATCAAC   | 3035 |
| QY   | 101    | AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys  | 120  |
| Db   | 3036   | AACAACCTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGTCTTTAACAATTCAAGTCAA | 3095 |
| QY   | 121    | GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln  | 140  |
| Db   | 3096   | GAGGTCAGCGAGAATGACGGTAGCAGCAGGATTGGCCAATAACCTTACCAGCACGGTTT   | 3155 |
| QY   | 141    | ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys  | 160  |
| Db   | 3156   | GTGTTTACTGACTCGGAGTACCAGCTCCCCTACGTGTCGGGTGGCGCACCAAGCTGT     | 3215 |
| QY   | 161    | LeuProPropheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn  | 180  |
| Db   | 3216   | CTCCGCGCGTTTCCAGCGGACGTCTTCATGTCCTCCCTAGTATGGATACCTCACCC      | 3275 |
| QY   | 181    | AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln  | 200  |
| Db   | 3276   | AACGGAAGTCAAGCGGTGGGACGCTCATCTTTTACTGCTGGAGTACTTCCCTTCGCAG    | 3335 |
| QY   | 201    | MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis  | 220  |
| Db   | 3336   | ATGCTAAGGACTGGAAATAACTTCCAATTTCAGCTATACCTTCGAGGATGTACCTTTT    | 3395 |
| QY   | 221    | SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr  | 240  |
| Db   | 3396   | AGCAGCTACGCTCACAGCCAGAGTTTGGATCGCTTGATGAATCCTCTTATTGATCAGTAT  | 3455 |
| QY   | 241    | LeuTyrTyrLeuAsnArgThrGln---AsnGlnSerGlySerAlaGlnAsnLysAspLeu  | 259  |
| Db   | 3456   | CTGTACTACCTGAACAGAACCGCAAGGAACAACCTCTGGAACAACCAACCAATCACGG    | 3515 |
| QY   | 260    | LeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGly  | 279  |
| Db   | 3516   | CTTTTATAGCCAGGCTGGGCCCTCAGTCTATGTCTTTGCGGCCAGAAATTTGGCTAC     | 3575 |
| QY   | 280    | ProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPhe  | 299  |
| Db   | 3576   | CCCTGTACCGGCAACAGAGACTTTCAAAGACTGCTAACGACAACAACAACAGTAAC      | 3635 |
| QY   | 300    | ThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGly  | 319  |
| Db   | 3636   | CCTTGGACAGCGCCAGCAAAATATCATCTCAATGGCCGCGACTCGTGGTGAATCCAG     | 3695 |
| QY   | 320    | ThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMet  | 339  |
| Db   | 3696   | CCAGCTATGGCCAGTCAACAGGACGATGAAGAAAAATTTTCCCTATGCACGGCAATCTA   | 3755 |

|   |           |   |      |
|---|-----------|---|------|
| QY  | 340       | IlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThr    | 359  |
| Db  | 3756      | ATATTTGGCAAGAAGGACAAACGGCAAGTAACGCAGAATTAGATAATGTAATGATTACG     | 3815 |
| QY  | 360       | AspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAla    | 379  |
| Db  | 3816      | GATGAAGAAGAGATTGGTACCACCAATCCTGTGGCAACAGAGCAGTATGGAACTGTGGCA    | 3875 |
| QY  | 380       | ValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAla    | 399  |
| Db  | 3876      | AATAACTTGCAGAGCTCAAATACAGCTCCACGACTGGAACCTGTCAATCATCAGGGGCC     | 3935 |
| QY  | 400       | LeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLys    | 419  |
| Db  | 3936      | TTACCTGGCATGGTGTGGCAAGATCGTGACGTGTACCTTCAAGGACCTATCTGGGCAAA     | 3995 |
| QY  | 420       | IleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLys    | 439  |
| Db  | 3996      | ATTCTTCACACCGGATGGACACTTTCATCCTCTCTCTCTGATGGAGGCTTTGGACTGAA     | 4055 |
| QY  | 440       | AsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGlu       | 459  |
| Db  | 4056      | CATCCGCCTCCTCAAATCATGATCAAAAATACTCCGGTACCTCGGAATCCTTCGACCACC    | 4115 |
| QY  | 460       | PheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerVal    | 479  |
| Db  | 4116      | TTCAGTGGCGGAAAGTTTGTCTTCTTCATCACACAGTACTCCACGGGACAGGTCAGCGTG    | 4175 |
| QY  | 480       | GluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyr    | 499  |
| Db  | 4176      | GAGATCGAGTGGGAGCTGCAGNAGGAAAAACAGCAAAACGCTGGAAATCCCGAAATTCAGTAC | 4235 |
| QY  | 500       | ThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyr    | 519  |
| Db  | 4236      | ACTTCCAACCTACAACAAGTCTGTTAATCGTGGACTTACCGTGGATACTAATGGCGTGTAT   | 4295 |
| QY  | 520       | ThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu                   | 534  |
| Db  | 4296      | TCAGAGCCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTG                   | 4340 |
| RESULT 7  |           |   |      |
| US-09-804-898-1   |           |   |      |
| ; Sequence 1, Application US/09804898                       |           |   |      |
| ; Patent No. US20020045264A1                                |           |   |      |
| ; GENERAL INFORMATION:                                      |           |   |      |
| ; APPLICANT: DURING, MATTHEW                                |           |   |      |
| ; APPLICANT: XIAO, WEIDONG                                  |           |   |      |
| ; TITLE OF INVENTION: PRODUCTION OF CHIMERIC CAPSID VECTORS |           |   |      |
| ; FILE REFERENCE: 102182-14                                 |           |   |      |
| ; CURRENT APPLICATION NUMBER: US/09/804,898                 |           |   |      |
| ; CURRENT FILING DATE: 2001-03-13                           |           |   |      |
| ; PRIOR APPLICATION NUMBER: 60/189,110                      |           |   |      |
| ; PRIOR FILING DATE: 2000-03-14                             |           |   |      |
| ; NUMBER OF SEQ ID NOS: 6                                   |           |   |      |
| ; SOFTWARE: PatentIn version 3.0                            |           |   |      |
| ; SEQ ID NO 1   |           |   |      |
| ; LENGTH: 4679  |           |   |      |
| ; TYPE: DNA   |           |   |      |
| ; ORGANISM: adeno-associated virus 2                        |           |   |      |
| US-09-804-898-1   |           |   |      |
| Alignment Scores:   |           |   |      |
| Pred. No.:  | 7.92e-279 | Length:   | 4679 |
| Score:  | 2481.50   | Matches:  | 445  |
| Percent Similarity:   | 90.26%    | Conservative:   | 37   |
| Best Local Similarity:                                      | 83.33%    | Mismatches:   | 51   |
| Query Match:  | 85.39%    | Indels:   | 1    |
| DB:   | 9         | Gaps:   | 1    |
| US-09-807-802A-17 (1-534) x US-09-804-898-1 (1-4679)        |           |   |      |
| QY  | 1         | MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly       | 20   |

Db 2809 ATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGGCGCGACGGAGTGGGT 2868

QY 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40

Db 2869 AATTCTCGGAAATTGGCATTGGCATTCACATGGATGGGCGACAGAGTCATCACCACC 2928

QY 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60

Db 2929 AGCACCCGAACCTGGGCGCTGCCACCTACAACAACCAACCTCTACAACAAATTTCCAGC 2988

QY 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80

Db 2989 CAATCA---GGAGCCTCGAACGACAATCACTATTGGCTACAGCACCCCTTGGGGGTAT 3045

QY 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100

Db 3046 TTTGACTTCAACAGATTCCACTGCCACTTTTCACCACCGTACGTGGCAAGACTCATCAAC 3105

QY 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120

Db 3106 AACAACTGGGATTCGACCCCAAGAGACTCAACTTCAAGCTCTTAAACATTCAAGTCAAA 3165

QY 121 GluValThrThrAsnAspGlyValThrIleAlaAsnAsnLeuThrSerThrValGln 140

Db 3166 GAGGTACCGCAGAATGACGGTACGACGACGATTGCCAATAACCTTACCAGCACGGTTTAC 3225

QY 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160

Db 3226 GTGTTTACTGACTCGGAGTACCGACTCCGCTACGTCCTCGGCTCGGCGCATCAAGGATGC 3285

QY 161 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180

Db 3286 CTCCCGCGGTTCGCCAGCAGACGCTCTTCATGGTGCCACAGTATGGATACCTCACCCCTGAAC 3345

QY 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200

Db 3346 AACGGGAGTCAGCGAGTAGGACGCTCTTCAATTTACTGCTGGAGTACTTTCCTTCTCAG 3405

QY 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPropheHis 220

Db 3406 ATGCTGCGTACCGGAAACAACCTTACCCTCAGCTACACTTTTGAGGACGTTCTCTTCCAC 3465

QY 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240

Db 3466 AGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGACCAGTAC 3525

QY 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260

Db 3526 CTGTATTACTTGAGCAGAAACAACACTCCAAGTGGAAACCAACCACGAGTCAAGGCTTCAG 3585

QY 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280

Db 3586 TTTTCTCAGGCCGAGCGAGTGACATTCCGGACCACTGCTAGGAAGTCTCTCTGGACCC 3645

QY 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300

Db 3646 TGTTACCGCCAGCAGCGAGTATCAAAGACATCTGCGGATAACAACAACAGTGAATACTCG 3705

QY 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320

Db 3706 TGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGTTGAATCCGGGCCCG 3765

QY 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340

Db 3766 GCCATGGCAAGCCACAAGCAGATGAAGAAAAGTTTTTCTCAGAGCGGGGTCTCATC 3825

QY 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360

Db 3826 TTTGGGAAGCAAGGCTCAGAGAAAACAATGTGGACATTGAAAGGTGATGATTACAGAC 3885

QY 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380

Db 3886 GAAGAGGAAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTATGGTTCTGTATCTACC 3945

QY 381 AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400

Db 3946 AACCTCCAGAGAGGCAACAGACAAGCAGCTACCGAGATGTCAACACACAAGCGCTTCTT 4005

QY 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420

Db 4006 CCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGCAAGATT 4065

QY 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440

Db 4066 CCACACACGGACGGACATTTTCACCCCTCTCCCTCATGGGTGGATTTCGACTTAAACAC 4125

QY 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460

Db 4126 CCTCCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGGAATCCTTCGACCACCTTC 4185

QY 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480

Db 4186 AGTGGCGCAAGTTTGCTTCCTTCATCACACAGTACTCCACGGGACAGGTCAGCGTGGAG 4245

QY 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 500

Db 4246 ATCGAGTGGAGCTGCAGAAGGAAAACAGCAACGCTGGAATCCCGAAATTCAGTACACT 4305

QY 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAsnAsnAsnGlyLeuTyrThr 520

Db 4306 TCCAACACTACAACAAGTCTGTTAATGTGGACTTTACTGTGGACACTAATGGCGTGTATTCA 4365

QY 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534

Db 4366 GAGCCTCGCCCCCATTTGGCACCAGATACCTGACTCGTAATCTG 4407

RESULT 8

US-09-945-681-10  
; Sequence 10, Application US/09945681  
; Patent No. US20020064878A1  
; GENERAL INFORMATION:  
; APPLICANT: UNIVERSITE DE NANTES  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR RECOMBINANT  
; TITLE OF INVENTION: ADENO-ASSOCIATED VIRUS PRODUCTION  
; FILE REFERENCE: B4182AA - UNIVERSITE DE NANTES  
; CURRENT APPLICATION NUMBER: US/09/945,681  
; CURRENT FILING DATE: 2001-09-05  
; PRIOR APPLICATION NUMBER: PCT/EP 00/01854  
; PRIOR FILING DATE: 2000-03-05  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 4679  
; TYPE: DNA  
; ORGANISM: adeno-associated virus 2  
US-09-945-681-10

Alignment Scores:  
Pred. No.: 7.92e-279 Length: 4679  
Score: 2481.50 Matches: 445  
Percent Similarity: 90.26% Conservative: 37  
Best Local Similarity: 83.33% Mismatches: 51  
Query Match: 85.39% Indels: 1  
DB: 9 Gaps: 1

US-09-807-802A-17 (1-534) x US-09-945-681-10 (1-4679)

QY 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20

Db 2809 ATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGGCGCGAGTGGGT 2868

QY 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40

Db 2869 AATTCTCGGAAATTGGCATTCGATTCCACATGGATGGGCGACAGAGTCATCACCACC 2928

QY 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60

Db 2929 AGCACCCGAACCTGGGCCCTGCCACCTACAAACACACCTCTACAAACAAATTTCCAGC 2988  
Qy 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80  
Db 2989 CAATCA---GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTGGGGGTAT 3045  
Qy 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTyrGlnArgLeuIleAsn 100  
Db 3046 TTTGACTTCAACAGATTCCACTGCCACTTTTCCACCAAGTACTGGCAAAAGACTCATCAAC 3105  
Qy 101 AsnAsnTyrGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120  
Db 3106 AACAACTGGGGATTCCGACCCCAAGACTCAACTTCAAGCTCTTTAACATTCAAGTCAA 3165  
Qy 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140  
Db 3166 GAGGTCACGCAGAAATGACGGTACGACGAGATTGCCAATAACCTTACCAGCAGGTTTCAG 3225  
Qy 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160  
Db 3226 GTGTTTACTGACTCGGAGTACCAGCTCCCGTACGCTCGGCTCGGCGCATCAAGGATGC 3285  
Qy 161 LeuProPropheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
Db 3286 CTCCCGCGTTTCCCAGCAGACGCTTTCATGGTGGCCAGATGGATACCTCACCTGAAC 3345  
Qy 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
Db 3346 AACGGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCCTGGAGTACTTTCTTCTCAG 3405  
Qy 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis 220  
Db 3406 ATGCTGCGTACCGGAACAACACTTTACCTTCAGCTACACTTTTGAGGACGTTCTTTCCAC 3465  
Qy 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
Db 3466 AGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGACCAGTAC 3525  
Qy 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260  
Db 3526 CTGTATTACTTGAGCAGAAACAACACTCCAAGTGGAAACCACCACGCGAGTCAAGCTTCAG 3585  
Qy 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280  
Db 3586 TTTTCTCAGGCGGAGCGAGTGACATTCGGGACCAGTCTAGGAACCTGGCTTCTCTGGACCC 3645  
Qy 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
Db 3646 TGTTACCGCCAGCAGCGAGTATCAAAAGACATCTCGGGAATAACAACAACAGTGAATACTCG 3705  
Qy 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320  
Db 3706 TGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGCCCG 3765  
Qy 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340  
Db 3766 GCCATGGCAAGCCACAAGGACGATGAAGAAAGTTTTTTCTCAGAGCGGGGTTCTCATC 3825  
Qy 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360  
Db 3826 TTTGGGAAGCAAGGCTCAGAGAAAACAATAATGTGGACATTGAAAAGGTTCATGATTACAGAC 3885  
Qy 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380  
Db 3886 GAAGAGGAAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTATGGTTCTGTATCTACC 3945  
Qy 381 AanPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400  
Db 3946 AACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAACACACAAGGCGTTCTT 4005  
Qy 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420

Db 4006 CCAGGCATGCTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGCAAGATT 4065  
Qy 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440  
Db 4066 CCACACACGAGCGGACATTTTCCACCCCTCTCCCTCATGGGTGGATTGGACTTAAACAC 4125  
Qy 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460  
Db 4126 CCTCCTCCACAGATTCTCATCAAGAAACACCCCGGTACCTCGGAATCCTTCGACCACCTTC 4185  
Qy 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480  
Db 4186 AGTCGGGCAAAGTTTGCTTCTTCATCACACAGTACTCCACGGGACAGGTGAGGTGGAG 4245  
Qy 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 500  
Db 4246 ATCGAGTGGGAGCTGCAGAAGGAAACAGCAACCGTGGAAATCCCGAAATTCAGTACACT 4305  
Qy 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520  
Db 4306 TCCAACACTACAACAAGTCTGTTAATGTGGACTTTACTGTGGACACTAATGGCGTGTATTCA 4365  
Qy 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
Db 4366 GAGCCTCGCCCCATTGGCACCAGATACCTGACTCGTAATCTG 4407  
RESULT 9  
US-10-136-819-6  
; Sequence 6, Application US/10136819  
; Publication No. US20030166593A1  
; GENERAL INFORMATION:  
; APPLICANT: Chien, Kenneth  
; APPLICANT: Hoshijima, Masahiko  
; TITLE OF INVENTION: No. US20030166593A1-viral vesicle vector for cardiac specific ge  
; FILE REFERENCE: 6627-PA1198  
; CURRENT APPLICATION NUMBER: US/10/136,819  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: 60/287,423  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 4679  
; TYPE: DNA  
; ORGANISM: adeno-associated virus 2  
US-10-136-819-6  
Alignment Scores:  
Pred. No.: 7.92e-279 Length: 4679  
Score: 2481.50 Matches: 445  
Percent Similarity: 90.26% Conservative: 37  
Best Local Similarity: 83.33% Mismatches: 51  
Query Match: 85.39% Indels: 1  
DB: 13 Gaps: 1  
US-09-807-802A-17 (1-534) x US-10-136-819-6 (1-4679)  
Qy 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20  
Db 2809 ATGGCTACAGCGAGTGGCGCACCAATGGCAGACAATAACGAGGGCGCCGACGAGTGGGT 2868  
Qy 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40  
Db 2869 AATTCTCGGGAATTTGGCATTTCCGATTCCCATGGATGGGCGACAGAGTCATCACCACC 2928  
Qy 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
Db 2929 AGCACCCGAACCTGGGCCCTGCCACCTACCAACCAACCACTCTACAAACAATTTCCAGC 2988  
Qy 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80  
Db 2989 CAATCA---GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTGGGGGTAT 3045



QY 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100  
Db 3046 TTTGACTTCAACAGATTCCCACTGCCACTTTTACCACCGTACTGGCAAGACTCATCAAC 3105  
QY 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120  
Db 3106 AACAACTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAAACATTCAAGTCAAA 3165  
QY 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140  
Db 3166 GAGGTACGCAGAAATGACGGTACGACGACGATTGCCAATAACTTACCAGCACGGTTTCAG 3225  
QY 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160  
Db 3226 GTGTTTACTGACTCGGAGTACCAAGCTCCCGTACGTCTCCGGCTCGGCGCATCAAGGATGC 3285  
QY 161 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
Db 3286 CTCCCGCGGTTCCAGCAGACAGCTTTTCATGGTGCACAGTATGGATACCTCACCCCTGAAC 3345  
QY 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
Db 3346 AACGGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCTCGAGTACTTTCTCTCTCAG 3405  
QY 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis 220  
Db 3406 ATGCTGCTACCGGAACAACACTTTTACTCTCAGCTACACTTTTGAGGACGTTCTCTTCCAC 3465  
QY 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
Db 3466 AGCAGCTACGCTCACAGCCAGAGCTCGACCGTCTCATGAATCCTCTCATCGACCAGTAC 3525  
QY 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260  
Db 3526 CTGTATTACTTGAGCAGAAACAACACTCCAAGTGGAAACCACCACGAGTCAAGGCTTCAG 3585  
QY 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280  
Db 3586 TTTTCTCAGGCGGAGCGAGTGACATTGCGGACCAGTCTAGGAAGTGGCTTCTCTGGACCC 3645  
QY 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
Db 3646 TGTTACCGCCAGCAGCGAGTATCAAAAGACATCTCGGATAACAACAACAGTGAATACTCG 3705  
QY 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320  
Db 3706 TGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGTGAATCCGGGCCCG 3765  
QY 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340  
Db 3766 GCCATGGCAAGCCACAAGCAGATGAAGAAAGTTTTTCTCTCAGAGCGGGGTTCTCATC 3825  
QY 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360  
Db 3826 TTTGGGAAGCAAGGCTCAGAGAAAACAATGTGGACATTGAAAGGTCAATGATTACAGAC 3885  
QY 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380  
Db 3886 GAAGAGGAAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTATGGTTCTGTATCTACC 3945  
QY 381 AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400  
Db 3946 AACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAACACACAAGGCGTTCTT 4005  
QY 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420  
Db 4006 CCAGGCATGGTCTGGCAGGACAGAGATGTGTACTTTCAGGGGCCCATCTGGGCAAGATT 4065  
QY 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440  
Db 4066 CCACACACGGACGGACATTTTCACCCCTCTCCCTCATGGTGGATTTCGACTTAAACAC 4125  
QY 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460

Db 4126 CCTCTCCACAGATTCTCATCAAGAACACCCCGTACCTCGGAATCCTTCGACCACCTTC 4185  
QY 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480  
Db 4186 AGTCGGGCAAGATTGCTTCTTCATCACACAGTACTCCACGGGACAGGTTCAGCTGGAG 4245  
QY 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 500  
Db 4246 ATCGAGTGGGAGCTGCAGAAGGAAACAGCAACCGTGAATCCCGAAATTTCAGTACACT 4305  
QY 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520  
Db 4306 TCCAACACTACAACAAGTCTGTTAATGTGGACTTTTACTGTGGACACTAATGGCGTGTATTCA 4365  
QY 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
Db 4366 GAGCCTCGCCCCATTGGCACCAGATACCTGACTCGTAATCTG 4407

RESULT 10  
US-10-038-972A-12  
; Sequence 12, Application US/10038972A  
; Publication No. US20020192823A1  
; GENERAL INFORMATION:  
; APPLICANT: J. Bartlett  
; TITLE OF INVENTION: AAV VECTORS AND METHODS  
; FILE REFERENCE: 28335/36996US  
; CURRENT APPLICATION NUMBER: US/10/038,972A  
; CURRENT FILING DATE: 2002-01-04  
; PRIOR APPLICATION NUMBER: US 60/260,124  
; PRIOR FILING DATE: 2001-01-05  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 12  
; LENGTH: 4679  
; TYPE: DNA  
; ORGANISM: adeno-associated virus 2  
US-10-038-972A-12

Alignment Scores:  
Pred. No.: 7.92e-279 Length: 4679  
Score: 2481.50 Matches: 445  
Percent Similarity: 90.26% Conservative: 37  
Best Local Similarity: 83.33% Mismatches: 51  
Query Match: 85.39% Indels: 1  
DB: 14 Gaps: 1

US-09-807-802A-17 (1-534) x US-10-038-972A-12 (1-4679)

QY 1 MetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20  
Db 2809 ATGGTACAGGCAGTGGCGCACCAATGGCAGACAAATAACGAGGGCGCCGCGGAGTGGGT 2868  
QY 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40  
Db 2869 AATTCTCGGAAATTGGCATTTGGCATTCGATTCCACATGGATGGGCGACAGAGTCATCACCAC 2928  
QY 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
Db 2929 AGCACCAGAACCTGGGCCCTGCCACCTACACACCACCTCTACAAACAATTTCCAGC 2988  
QY 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80  
Db 2989 CAATCA---GGAGCCTCGAACGACAATCACTACTTTTGGTTACAGCACCCCTTGGGGGTAT 3045  
QY 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100  
Db 3046 TTTGACTTCAACAGATTCCACTGCCACTTTTTCACCACCGTACTGGCAAGAGCTCATCAAC 3105  
QY 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120  
Db 3106 AACAACTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAAACATTCAAGTCAAA 3165

Qy 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140  
Db 3166 GAGGTACGCAGAATGACCGGTACGACGAGATTGCCAATAACCTTACCAGCACGGTTCAG 3225  
Qy 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160  
Db 3226 GTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTCCTCGGCTCGGCGCATCAAGGATGC 3285  
Qy 161 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
Db 3286 CTCCGCGGTTCACAGCAGACGCTTTCATGGTGTCCACAGTATGGATACCTCACCCCTGAAC 3345  
Qy 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
Db 3346 AACGGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCTGGAGTACTTTCTCTTCAG 3405  
Qy 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis 220  
Db 3406 ATGCTGCGTACCGGAACAACATTACCTTCAGCTACACTTTTGAGGACGTTCTCTTCCAC 3465  
Qy 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
Db 3466 AGCAGTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGACCGATAC 3525  
Qy 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260  
Db 3526 CTGTATTACTTGAGCAGACAACAACACTCCAAGTGGAAACCACACGACGAGTCAAGGCTTCAG 3585  
Qy 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTyrLeuProGlyPro 280  
Db 3586 TTTTCTCAGCGCGGAGCGAGTGACATTCGGGACCAGTCTAGGAACGTGGCTTCCTGGACCC 3645  
Qy 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
Db 3646 TGTTACCGCCAGCAGCGAGTATCAAGACATCTCGCGATAACAACAACAGTGAATACTCG 3705  
Qy 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320  
Db 3706 TGGACTGGAGTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGCGCG 3765  
Qy 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340  
Db 3766 GCCATGGCAAGCCACAAGGACGATGAAGAAAGTTTTTTTCTCAGAGCGGGTTCTCATC 3825  
Qy 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360  
Db 3826 TTTGGGAAGCAAGGCTCAGAGAAACAARAATGTGGACATTGAAGAAGGTCTATGATTACAGAC 3885  
Qy 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380  
Db 3886 GAAGAGGAATATCAGGACAACCAATCCGTGGCTACGGAGCAGTATGGTTCTGTATCTACC 3945  
Qy 381 AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400  
Db 3946 AACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAACACACACAAGGCGTCTT 4005  
Qy 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420  
Db 4006 CCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGCAAGATT 4065  
Qy 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440  
Db 4066 CCACACACGGACGGACATTTTTCACCCCTCTCCCTCATGGGTGGATTCCGACTTAAACAC 4125  
Qy 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460  
Db 4126 CCTCCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAATCCTTCGACCACCTTC 4185  
Qy 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480  
Db 4186 AGTGGGCAAAAGTTTGCTTCTTCATCACACAGTACTCCACGGGACAGGTACGGGTGGAG 4245  
Qy 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 500

Db 4246 ATCGAGTGGAGCTGCAGAAGGAAACAGCAAAACGCTGGAATCCCGAATTCAGTACACT 4305  
Qy 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520  
Db 4306 TCCAACCTACAACAAGTCTGTTAATGTGGACTTTACTGTGGACACTAATGGCGGTATTCA 4365  
Qy 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
Db 4366 GAGCCTCGCCCCCATTTGGCACCAGATACCTGACTCGTAATCTG 4407  
RESULT 11  
US-10-291-583-120  
; Sequence 120, Application US/10291583  
; Publication No. US20030138772A1  
; GENERAL INFORMATION:  
; APPLICANT: Gao, Guangping  
; APPLICANT: Wilson, James M.  
; APPLICANT: Alvira, Mauricio  
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus  
; FILE REFERENCE: UPN-02735USA  
; CURRENT APPLICATION NUMBER: US/10/291,583  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: US 60/350,607  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/341,117  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: US 60/377,066  
; PRIOR FILING DATE: 2002-05-01  
; PRIOR APPLICATION NUMBER: US 60/386,675  
; PRIOR FILING DATE: 2002-06-05  
; NUMBER OF SEQ ID NOS: 120  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 120  
; LENGTH: 2205  
; TYPE: DNA  
; ORGANISM: adeno-associated virus serotype, clone A3.1vp1  
US-10-291-583-120  
Alignment Scores:  
Pred. No.: 1.52e-276 Length: 2205  
Score: 2457.50 Matches: 442  
Percent Similarity: 89.70% Conservative: 37  
Best Local Similarity: 82.77% Mismatches: 54  
Query Match: 84.57% Indels: 1  
DB: 13 Gaps: 1  
US-09-807-802A-17 (1-534) x US-10-291-583-120 (1-2205)  
Qy 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnGluGlyAlaAspGlyValGly 20  
Db 607 ATGGCTTCAGCGGTGGGCACCAATGGCAGACAATAACGAAGCGCGCAGGAGTGGGT 666  
Qy 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40  
Db 667 AATTCTCGGGAATTGGCATTTCCGATTCCACATGGATGGCGCAGAGTTATCACCACC 726  
Qy 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
Db 727 AGCACAAGAACCTGGGCCCTCCCCACCTACAATAATCACCTCTACAAGCAATCTCCAGC 786  
Qy 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80  
Db 787 GAATCG--GGAGCCACCAACGACAACCACTACTTCGGCTACAGCACCCCTGGGGGTAT 843  
Qy 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100  
Db 844 TTTGACTTTAACAGATTCCACTGTCACTTCTCACACGACTGTCGACGCGACTCATCAAC 903  
Qy 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120  
Db 904 AACAACTGGGGATTAGACCCCAAGAACTCAATTCAAGCTCTTCAACATCCCAAGTCAAG 963

Qy 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140  
Db 964 GAGGTCACGCAGAATGATGGAACCCAGACATCGCCAATAACCTTACCAGCACGGTGCG 1023  
Qy 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160  
Db 1024 GTCTTCACAGACTCTGAGTACCAGCTGCCCTACGCTCTCGGTTCCGCTCACCAGGGCTGC 1083  
Qy 161 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
Db 1084 CTTCCGCCGTTCCAGCAGACGCTCTTCATGATTCCTCAGTACGGCTACTTGACTCTGAAC 1143  
Qy 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
Db 1144 AATGGCAGCCAAAGCGGTAGGACGTTCTTCATCTACTGTCTAGAGTATTTTCCCTCTCAG 1203  
Qy 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPropheHis 220  
Db 1204 ATGCTGAGGACGGGAACAACCTTCACCTTCAGCTACACTTTTGAAGACGTGCCTTTCCAC 1263  
Qy 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
Db 1264 AGCAGCTACGCGACAGCCAGAGTCTGGATCGGCTGATGAATCCTCTCATTTGACCAATAC 1323  
Qy 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260  
Db 1324 CTGTATTACCTGAGCAAAACTCAGGGTACAAGTGAACAACCGCAGCAATCGAGACTGCAG 1383  
Qy 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280  
Db 1384 TTCAGCCAAGCTGGGCCTAGCTCCATGGCTCAGCAGGCCAAAACTGGCTACCGGACCC 1443  
Qy 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
Db 1444 AGCTACCGACAGCAGCGAATGTCTAAGACGGCTAATGACAACAACACAGTGAATTTGCT 1503  
Qy 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320  
Db 1504 TGGACTGCAGCCACCAAAATATTACCTGAATGGAAGAAATTTCTCTGGTCAATCCCGGCC 1563  
Qy 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340  
Db 1564 CCAATGGCCAGTCACAAGGACGATGAGGAAAAGTATTTCCCATGACGGAAATCTCATC 1623  
Qy 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360  
Db 1624 TTTGGAAAACAAGGCACAGGAACACTACCAATGTGGACATTTGAATCAGTGCTATTACAGAC 1683  
Qy 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380  
Db 1684 GAAGAAGAAATCAGAAACAATAATCCTGTGGCTACAGAACNATACGGACAGGTTGCCACC 1743  
Qy 381 AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400  
Db 1744 AACCATCAGAGTCAGAACACCACAGCTTCCATGGAAGTGTGGACAGCCAGGGAATCTTA 1803  
Qy 401 ProGlyMetValTrpGlnAspArgAspValTyrleuGlnGlyProIleTrpAlaLysIle 420  
Db 1804 CCTGGAATGGTGTGGCAGGACCGCGATGTCTATCTTCAAGGTCCCATTGGGCCAAAAC 1863  
Qy 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440  
Db 1864 CCTCACCGGACGGACACTTTTCATCCTTCCGCTCATGGAGGCTTTGGACTGAAACAC 1923  
Qy 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460  
Db 1924 CCTCCTCCCCAGATCCTGATCAAAAACACACCTGTGCCAGCGAATCCCGGACCACTTTC 1983  
Qy 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480  
Db 1984 ACTCCTGGAAAGTTTGCTTCGTTTATTACCCAGTATTCCACCGGACAGGTCAGCGTGGAA 2043

Qy 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 500  
Db 2044 ATAGAGTGGGAGCTGCAGAAAGAAAACAGCAACCGCTGGAACCCAGAAATTCAGTACACC 2103  
Qy 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520  
Db 2104 TCCAACACTACAACAGTCGGTGAATGTGGAGTTTACCGTGGACGCAACGGTGTATTCT 2163  
Qy 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
Db 2164 GAACCCCGCCCTATTGGCACTCGTTACCTTACCCCGGAACCTTG 2205

RESULT 12  
US-10-291-583-54  
; Sequence 54, Application US/10291583  
; Publication No. US20030138772A1  
; GENERAL INFORMATION:  
; APPLICANT: Gao, Guangping  
; APPLICANT: Wilson, James M.  
; APPLICANT: Alvira, Mauricio  
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus  
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1  
; FILE REFERENCE: UPN-02735USA  
; CURRENT APPLICATION NUMBER: US/10/291,583  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: US 60/350,607  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/341,117  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: US 60/377,066  
; PRIOR FILING DATE: 2002-05-01  
; PRIOR APPLICATION NUMBER: US 60/386,675  
; PRIOR FILING DATE: 2002-06-05  
; NUMBER OF SEQ ID NOS: 120  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 54  
; LENGTH: 3123  
; TYPE: DNA  
; ORGANISM: new AAV serotype, clone A3.4  
US-10-291-583-54

Alignment Scores:  
Pred. No.: 1.34e-275 Length: 3123  
Score: 2451.50 Matches: 441  
Percent Similarity: 89.51% Conservative: 37  
Best Local Similarity: 82.58% Mismatches: 55  
Query Match: 84.36% Indels: 1  
DB: 13 Gaps: 1

US-09-807-802A-17 (1-534) x US-10-291-583-54 (1-3123)

Qy 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20  
Db 1444 ATGGCTTCAGGCGGTGGGGACCAATGGCAGACGATAACGAAGCGCCGACGGAGTGGGT 1503  
Qy 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40  
Db 1504 AATCTCTCGGAAATTTGGCATTTCCGATTTCCACATGGATGGGCGACAGAGTTATCACCACC 1563  
Qy 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
Db 1564 AGCAAGAACCTGGGCCCCCTCCACCTACAATAATCACCTCTACAAGCAAAATCTCCAGC 1623  
Qy 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80  
Db 1624 GAATCG--GGAGCCACCACGACAACCACTACTTCGGCTACAGCACCCCTCGGGGTAT 1680  
Qy 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100  
Db 1681 TTTGACTTTAACAGATTCCACTGTCTACTTCTCACCACGTGACTGGCAGCGACTCATCAAC 1740  
Qy 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnVallys 120



Db 1741 AACAACTGGGGATTAGACCCCAAGAAACTCAATTTCAAGCTCTTCAACATCCAAGTCAAG 1800

Qy 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140

Db 1801 GAGGTCACGCAGAAATGATGGAAACACGACCATCGCCAATAACCTTACCAGACGGTGCAG 1860

Qy 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160

Db 1861 GTCTTCACAGACTCTGAGTACCAGCTGCCCTACGTCCTCGGTTCCGGCTCACCGGGCTGC 1920

Qy 161 LeuProPropheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180

Db 1921 CTTCCGCCGTTCCACAGCAGACGCTCTTCATGATTCTCAGTACGGCTACTTGACTCTGAAC 1980

Qy 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200

Db 1981 AATGGCAGCCAAGCGGTAGACGTTCTTCATTCTACTGTCTAGAGTATTTCCCTCTCAG 2040

Qy 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPropheHis 220

Db 2041 ATGCTGAGGACGGGAAACAACCTTCACCTTCAGCTACACTTTTGAAGACGTGCTTCCAC 2100

Qy 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240

Db 2101 AGCAGCTACGGGCACAGCCAGAGTCTGGATCGGCTGATGAATCCTCTCATTGACCAGTAC 2160

Qy 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260

Db 2161 CTGTATTACCTGAGCAAAACTCAGGGTACAAGTGAACAACGCAGCAATCGAGACTGCAG 2220

Qy 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280

Db 2221 TTCAGCCAAGCTGGGCCTAGCTCCATGGCTCAGCAGGCCAAAACTGGCTACCGGGACCC 2280

Qy 281 CysTyrArgGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300

Db 2281 AGCTACCGACAGCGGAATGTCTAAGACGGCTAATGACAACAACAACAGTGAATTGCT 2340

Qy 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320

Db 2341 TGGACTGCAGCCACCACAATATTACCTGAATGGGAAGAAATCTCTGGTCAATCCGGGCCC 2400

Qy 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340

Db 2401 CCAATGGCCAGTCAAGAGCAGATGAGGAAAAGTATTTCCCATGCACGGAAATCTCATC 2460

Qy 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360

Db 2461 TTTGGAAAAACAAGGCACAGGAACCTACCAATGTGGACATGAATCAGTGTATTATACAGAC 2520

Qy 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380

Db 2521 GAAGAAGAAATCAGAAACAACCTAATCCTGTGGCTACAGAACAATACGGACAGGTGCCACC 2580

Qy 381 AsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400

Db 2581 AACCATCAGAGTCAGGACACACAGCTTCCTATGGAAGTGTGCACAGCCAGGGAATCTTA 2640

Qy 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420

Db 2641 CTTGGAATGGTGTGGCAGGACCGGATGCTATCTTCAAGGTCCCATTTTGGGCCAAAAC 2700

Qy 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440

Db 2701 CCTCACAGGACGGACACTTTTCATCTCTCCGCTCATGGAGGCTTTGGACTGAAACAC 2760

Qy 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460

Db 2761 CTTCTCCCCAGATCCTGTATCAAAAAACACACCTGTGCCAGCGAATCCCGGACCACTTTC 2820

Qy 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480

Db 2821 ACTCCTGGAAAGTTTGCTTCGTTTCATTACCCAGTATTCCACCGGACAGGTTCAGCGTGAA 2880

Qy 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 500

Db 2881 ATAGAGTGGGAGCTGCAGAAAGAAAACAGCAACGCTGGAACCCAGAAATTCAGTACACC 2940

Qy 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520

Db 2941 TCCTCACTACAACAAGTCGGTGAATGTGGAGTTTACCGTGGACGCAAAACGGTGTATTCT 3000

Qy 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534

Db 3001 GAACCCCGCCCTATTGGCACTCGTTACCTTACCCGGAACTTG 3042

RESULT 13

US-10-291-583-56

; Sequence 56, Application US/10291583

; Publication No. US20030138772A1

; GENERAL INFORMATION:

; APPLICANT: Gao, Guangping

; APPLICANT: Wilson, James M.

; APPLICANT: Alvira, Mauricio

; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus

; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1el Sequences Identifi

; FILE REFERENCE: UPN-02735USA

; CURRENT APPLICATION NUMBER: US/10/291,583

; CURRENT FILING DATE: 2002-11-12

; PRIOR APPLICATION NUMBER: US 60/350,607

; PRIOR FILING DATE: 2001-11-13

; PRIOR APPLICATION NUMBER: US 60/341,117

; PRIOR FILING DATE: 2001-12-17

; PRIOR APPLICATION NUMBER: US 60/377,066

; PRIOR FILING DATE: 2002-05-01

; PRIOR APPLICATION NUMBER: US 60/386,675

; PRIOR FILING DATE: 2002-06-05

; NUMBER OF SEQ ID NOS: 120

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 56

; LENGTH: 3122

; TYPE: DNA

; ORGANISM: new AAV serotype, clone A3.7

US-10-291-583-56

Alignment Scores:

Pred. No.: 2,29e-275 Length: 3122

Score: 2449.50 Matches: 441

Percent Similarity: 89.51% Conservative: 37

Best Local Similarity: 82.58% Mismatches: 55

Query Match: 84.29% Indels: 1

DB: 13 Gaps: 1

US-09-807-802A-17 (1-534) x US-10-291-583-56 (1-3122)

Qy 1 MetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20

Db 1453 ATGGCTTCAGCGGTGGGCACCAATGGCAGACAATAACGAAGCGCGACGGAGTGGGT 1512

Qy 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40

Db 1513 AATTCTCTCGGAAATTTGGCATTTCCGATTCACATGGATGGCGACAGAGTTATCACCACC 1572

Qy 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60

Db 1573 AGCACAAGAACCTGGGCCCTCCCCACCTACATAATATCGCCTCTACAGCAAAATCTCCAGC 1632

Qy 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80

Db 1633 GAATCG---GGAGCCACCAACGACAACCACTACTTCGGCTACAGCACCCCTGGGGGTAT 1689

Qy 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100

Db 1690 TTTGACTTTAACAGATTCCCACTGTCACTTCTCACCACTGACTGGCAGCGACTCATCAAC 1749

Qy 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120

Db 1750 AACAACTGGGGATTTAGACCCAAAGAACTCAATTTCAAGCTCTTCAACATCCAAGTCAAG 1809  
QY 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140  
Db 1810 GAGGTACGCAGAATGATGGAAACCAACCATCGCCAAATAACCTTACCAGCACGGTGCAG 1869  
QY 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160  
Db 1870 GTCTTCACAGACTCTGAGTACCAGCTGCGCCTACGTCCTCGGTTCCGCTCACCAGGGCTGC 1929  
QY 161 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
Db 1930 CTTCCGCCGTTCCACAGCAGACGCTCTTCATGATTCCTCAGTACGGTACTTGACTCTGAAC 1989  
QY 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
Db 1990 AATGGCAGCCCAAGCGGTAGGACGTTCTTCATCTACTGTCTAGAGTATTTTCCCTCTCAG 2049  
QY 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis 220  
Db 2050 ATGCTGAGGACGGGAACAACCTTCACCTTCAGCTACACTTTTGAAGACGTGCCTTTCCAC 2109  
QY 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
Db 2110 AGCAGCTACGGCACAGCCAGAGTCTGGATCGGCTGATGAATCCTCTCATTGACCAGTAC 2169  
QY 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260  
Db 2170 CTGTATTACCTGAGCAAAACTCAGGTFACAAGTGGAAACAACGACGAAATCGAGACTGCAG 2229  
QY 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280  
Db 2230 TTCAGCCAAGCTGGCCTAGCTCCATGGCTCAGCAGGCCAAAACTGGCTACCGGGACCC 2289  
QY 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
Db 2290 AGCTACCGCAGCAGCGAATGTCTAAGACGGTAAATGACAACAACAACAGTGAATTTGCT 2349  
QY 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleAsnProGlyThr 320  
Db 2350 TGGACTGCAGCCCAACAATATTACCTGAATGGAGAATAATTCTCTGGTCAATCCCGGGCCC 2409  
QY 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340  
Db 2410 CCAATGGCCAGTCACAAGGACGATGAGGAAAGTATTTCCCATGACGCGAAATCTCATC 2469  
QY 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360  
Db 2470 TTTGGAANAACAGGCACAGCAACTACCAATGTGGACATTTGAATCAGTGTATTACAGAC 2529  
QY 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380  
Db 2530 GAAGAAGAAATCAGAAACAACATAATCCTGTGGCTACAGAACNAATACGGACAGGTGGCCACC 2589  
QY 381 AsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400  
Db 2590 AACCATCAGAGTCAGAACACCACAGCTTCCTATGGAAGTGTGGACAGCCAGGGAATCTTA 2649  
QY 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420  
Db 2650 CCTGGAATGGTGTGGCAGGACCGGATGTCTATCTTCAAGGTCCCATTTGGGCCAAACT 2709  
QY 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440  
Db 2710 CCTCACACGGACGGACACTTTTCATCCTTCTCCGCTCATGGGAGGCTTTGGACTGAAACAC 2769  
QY 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460  
Db 2770 CCTCCTCCCCAGATCCTGATCAAAAACACACCTGTGCCAGCGAATCCCGGACCACCTTTC 2829  
QY 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480  
::: |||||

Db 2830 ACTCCTGGAAAGTTTGCTTCGTTTCATTACCACGATATTCCACGGACAGGTACCGTGGAA 2889  
QY 481 ileGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 500  
Db 2890 ATAGAGTGGAGCTGCAGAAAAGAAAACAGCAACGCTGGAAACCCAGAAATTCAGTACACC 2949  
QY 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520  
Db 2950 TCCAACACTACAACAGTCGGTGAATGTGGAGTTTACCGTGGACGCAAAACGGTGTATTCT 3009  
QY 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
Db 3010 GAACCCCGCCCTATTGGCACTCGTTACCTTACCTTACCCGGAACCTTG 3051  
RESULT 14  
US-10-291-583-55  
; Sequence 55, Application US/10291583  
; Publication No. US20030138772A1  
; GENERAL INFORMATION:  
; APPLICANT: Gao, Guangping  
; APPLICANT: Wilson, James M.  
; APPLICANT: Alvira, Mauricio  
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus  
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1el Sequences Identif  
; FILE REFERENCE: UPN-02735USA  
; CURRENT APPLICATION NUMBER: US/10/291,583  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: US 60/350,607  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/341,117  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: US 60/377,066  
; PRIOR FILING DATE: 2002-05-01  
; PRIOR APPLICATION NUMBER: US 60/386,675  
; PRIOR FILING DATE: 2002-06-05  
; NUMBER OF SEQ ID NOS: 120  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 55  
; LENGTH: 3113  
; TYPE: DNA  
; ORGANISM: new AAV serotype, clone A3.5  
US-10-291-583-55  
Alignment Scores:  
Pred. No.: 6.7e-275 Length: 3113  
Score: 2445.50 Matches: 440  
Percent Similarity: 89.51% Conservative: 38  
Best Local Similarity: 82.40% Mismatches: 55  
Query Match: 84.15% Indels: 1  
DB: 13 Gaps: 1  
US-09-807-802A-17 (1-534) x US-10-291-583-55 (1-3113)  
QY 1 MetaLaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20  
Db 1444 ATGGCTTCAGGCGGTGGGCGCAACCAATGGCAGACAATAACGAAAGCGCCGACGGAGTGGGT 1503  
QY 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40  
Db 1504 AATTCTCTCGGAAATTTGGCATTCGATTCACATGATGGATGGGCGACAGAGTTATCACCACC 1563  
QY 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
Db 1564 AGCACAGAACCTGGGCCCTCCACACCTACAATAATCACCTCTACAGCAATCTCCAGC 1623  
QY 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80  
Db 1624 GAATCG--GGAGCCACCAACGACACCACTACTTTCGGCTACAGACCCCTGGGGGTAT 1680  
QY 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100  
Db 1681 TTTGACTTTAAACAGATTCCCACTGTCTCCTCTCACCAACGTCGACTGGCAGCGACTCATCAAT 1740

Qy 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120  
Dd 1741 AACAACTGGGGATTAGACCCAGAAACTCAATTTCAAGCTCTTCAACATCCAGTCAAG 1800  
Qy 121 GluValThrThrAsnAspGlyValThrIleAlaAsnAsnLeuThrSerThrValGln 140  
Dd 1801 GAGGTACGCAGAAATGATGGAACACAGCACCATCGCCAAATAAATACCTACCAGCACGGTGCAG 1860  
Qy 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160  
Dd 1861 GTCTTCACAGACTCTGAGTACCAGCTGCCCTACGTCTCGTTCGGTCCAGGGGCTGC 1920  
Qy 161 LeuProPropheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
Dd 1921 CTTCCGCCGTTCCAGCAGACGCTTCTCATGATTCCTCAGTACGGCTACTTGACTCTGAAC 1980  
Qy 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
Dd 1981 AATGGCAGCCAAAGCGGTAGGACGTTCTTCTCATGATTCCTCAGTACGGCTACTTGACTCTGAAC 2040  
Qy 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPropheHis 220  
Dd 2041 ATGCTGAGGACGGGAAACAACTTCACTTCACTCAGCTACACTTTTGAAGACGTGCCTTTCCAC 2100  
Qy 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
Dd 2101 AGCAGCTACGCGCACAGCCAGAGTCTGGATCGGCTGATGAATCTCTCATTTGACCAGTAC 2160  
Qy 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260  
Dd 2161 CTGTATTACCTGAGCAAACTCAGGGTACAAGTGAACAACGACGCAATCGAGACTGCAG 2220  
Qy 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280  
Dd 2221 TTCAACCAAGCTGGCCTAGCTCCATGGCTCAGCAGGGCCAAACTGGTACCGGGGACCC 2280  
Qy 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
Dd 2281 AGCTACCGACAGCAGCGGAATGTCTAAGACGGCTAATGACAACAACAGTGAATTTGCT 2340  
Qy 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleAsnProGlyThr 320  
Dd 2341 TGGACTGCAGCCACCAATATTACCCGAATGGAAGAAATTTCTGTGTCAATCCCGGGCCC 2400  
Qy 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetile 340  
Dd 2401 CCAATGGCCAGTCACAAGGACGATGAGGAAAAAGTATTTCCCCATGCACGGAATCTCATC 2460  
Qy 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360  
Dd 2461 TTTGGAAAACAAGGCACAGGAACCTACCAATGTGGACATTGAATCAGTGTATTACAGAC 2520  
Qy 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380  
Dd 2521 GAAGAAGAAATCAGAACGACTAATCCTGTGGCTACAGAACAAATACGACAGGTTGCCACC 2580  
Qy 381 AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400  
Dd 2581 AACCGTCAGAGTCAGAAACACCACAGCTTCCATGGAAGTGTGGACAGCCAGGGAATCTTA 2640  
Qy 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420  
Dd 2641 CCTGGAATGGTGGCAGGACCGCGATGTCTATCTTCAAGGTCCCATTTGGGCCAAAAC 2700  
Qy 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440  
Dd 2701 CCTCACACGGACGGACACTTTTCATCTCTCCGCTCATGGGAGGCTTTGGACTGAAACAC 2760  
Qy 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460  
Dd 2761 CCTCTCCCCAGATCCTGTATCAAAAACACACCTGTGCCAGCGAATCCCGGACCACTTTC 2820  
Qy 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480

Dd 2821 ACTCCTGGAAAGTTTGCCTTCGTTTCATTACCCAGTATTCACCGACAGSTCAGCGTGGAA 2880  
Qy 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 500  
Dd 2881 ATAGAGTGGGAGCTGCAGAAAGAAAACAGCAACCGCTGGAACCCGGAAATTCAGTACACC 2940  
Qy 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520  
Dd 2941 TCCAACACTACAACAGTCGGTGAATGTGGAGTTTACCGTGGACGCAACCGTGTATTCT 3000  
Qy 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
Dd 3001 GAACCCCGCCCTATTGGCACTCGTTACCTTACCCCGAACTTG 3042  
RESULT 15  
US-10-291-583-35  
; Sequence 35, Application US/10291583  
; Publication No. US20030138772A1  
; GENERAL INFORMATION:  
; APPLICANT: Gao, Guangping  
; APPLICANT: Wilson, James M.  
; APPLICANT: Alvira, Mauricio  
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus  
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1el Sequences Identif  
; FILE REFERENCE: UPN-02735USA  
; CURRENT APPLICATION NUMBER: US/10/291,583  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: US 60/350,607  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/341,117  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: US 60/377,066  
; PRIOR FILING DATE: 2002-05-01  
; PRIOR APPLICATION NUMBER: US 60/386,675  
; PRIOR FILING DATE: 2002-06-05  
; NUMBER OF SEQ ID NOS: 120  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 35  
; LENGTH: 2489  
; TYPE: DNA  
; ORGANISM: new AAV serotype, clone 42.10  
US-10-291-583-35  
Alignment Scores:  
Pred. No.: 1.57e-274 Length: 2489  
Score: 2441.00 Matches: 443  
Percent Similarity: 90.64% Conservative: 41  
Best Local Similarity: 82.96% Mismatches: 46  
Query Match: 84.00% Indels: 4  
DB: 13 Gaps: 3  
US-09-807-802A-17 (1-534) x US-10-291-583-35 (1-2489)  
Qy 1 MetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20  
Dd 835 ATGGCTGCAGGGCGGTGGCGCTCCAATGGCAGACAATAACGAAGCGCCGAGTGGGT 894  
Qy 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40  
Dd 895 AATGCCTCCGAAATTTGGCAATTCGATTCACATGGCTGGCGCAGACAGAGTCATCACCACC 954  
Qy 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
Dd 955 AGCACCCGCACCTGGGCCCTGCCACCTACAAACCAACCACTCTTACAGCAGATATCAAGT 1014  
Qy 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80  
Dd 1015 CAGAGC---GGGGCTACCACGACAACCACTTCTTCGGTTCAGACACCCCTGGGGCTAT 1071  
Qy 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100  
Dd 1072 TTTGACTTCAACAGATTCCACTGCCACTTCTTCACCACTGGACTGGCAGCGACTCATCAAC 1131



```
QY 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPhelYsLeuPheAsnIleGlnVallys 120
Db 1132 AACAACTGGGGATTCCGGCCAGAAAGCTGCGGTTCAAGTTGTTCAACATCCAGGTCAAG 1191
QY 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140
Db 1192 GAGTCAACGACGACGACGCGGTTACGACCATCGCCATAAACCTTACCAGCACGATTGAT 1251
QY 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160
Db 1252 GTCTTCTCGGACTCGGAGTACCAACTGCCGTACGTCCTCGGCTCTGCGCACCGGCTGC 1311
QY 161 LeuProPropheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180
Db 1312 CTCCCTCCGTTCCCTGCGGACGTGTTTCATGATTCCTTCAGTACGGATATCTGACTCTAAC 1371
QY 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200
Db 1372 AACGGCAGTCAGTCTGTGGGACGTTCCTCTCTTACTGCTGGAGTACTTTTCTCTTCAG 1431
QY 201 MetLeuArgThrGlyAsnAsnPhelThrPheSerTyrThrPheGluGluValProPheHis 220
Db 1432 ATGCTGAGAAACGGGCAATAACTTTGAATTCAGCTACACCTTTGAGGAAGTGCTTTCCAC 1491
QY 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240
Db 1492 AGCAGCTATGCGCACAGCCAGAGCCTGGACCGGCTGATGAATCCCTCATCGACCATAC 1551
QY 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260
Db 1552 CTGTACTACCTGGCCCGGACCCAGAGCACTACGGGGTCC-----ACAAGGGAGCTGCAG 1605
QY 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280
Db 1606 TTCCATCAGGCTGGGCCCAACACCATGGCCGAGCAATCAAGAAGTGGTGGCCCGGACCC 1665
QY 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300
Db 1666 TGTTATCGGCAGCAGAGACTGTCAAAAACATAGACAGCAACAACAACAGTAACCTTTGCC 1725
QY 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320
Db 1726 TGGACCGGGGCCACTAAATACCATCTGAATGGTAGAAATTCATTAAACCAACCCGGGCGTA 1785
QY 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340
Db 1786 GCCATGGCCACCAACAAGGACGACGAGGACCACTTCTTCCCATCAACGGAGTGCTGGTT 1845
QY 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360
Db 1846 TTTGGCAAAACGGGGGCT--GCCAACAAGACACACGCTGGAAACGTCGTAATGACCAGC 1902
QY 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380
Db 1903 GAGGAGGAGATCAAAACCAACCAATCCCGTGGCTACAGAGATACGGTGTGGTCTCCAGC 1962
QY 381 AsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400
Db 1963 AACCTGCAATCGTCTACGGCCGACCCCGACACAGACTGTCAACAGCCAGGGGGCTCTG 2022
QY 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420
Db 2023 CCCGGCATGGTCTGGCAGAACCGGGACGTGTACCTGCAGGGTCCCATCTGGGCCAAAATT 2082
QY 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440
Db 2083 CCTCACACGGACGGCAACTTTCACCCGTCCTCCCTGATGGCGGATTTGGACTCAAAACAC 2142
QY 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460
Db 2143 CCGCCTCCTCAATTTCTCATCAAAAACACCCCGGTACCTGTCTAATCCTCCAGAGGTGTTT 2202
```

```
QY 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480
Db 2203 ACTCCTGCCAAAGTTTGCTTCATTTATCACGCAGTACAGCACCCGCCAGGTGAG 2262
QY 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTyrAsnProGluValGlnTyrThr 500
Db 2263 ATCGAGTGGGAAGTGCAGAAAGAAAACAGCAAAACGCTGGAATCCAGAGATTCAGTACAC 2322
QY 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520
Db 2323 TCAAAATTATGCCAAGTCTAATAATGTGAATTTGCTGTCAACAACGAAGGGTTTATACT 2382
QY 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534
Db 2383 GAGCCTCGCCCCCATTTGGCACCCGTTACTCCTCACCCGTAACCTG 2424
```

Search completed: February 15, 2004, 02:07:54  
Job time : 590.857 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 14, 2004, 15:19:20 ; Search time 7481.29 Seconds  
(without alignments)  
4024.639 Million cell updates/sec

Title: US-09-807-802A-13  
Perfect score: 3989  
Sequence: 1 MAADGYLPDWLEDNLSEGIR.....NNGLYTERPIGTRYLTRPL 736

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO spoel/US09807802/runat\_11022004\_175607\_15929/app\_query.fasta\_1.2389  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09807802@cgn\_1\_1\_12216@runat\_11022004\_175607\_15929 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

| Result No. | Score  | Query Match % | Length | DB | ID       | Description        |
|------------|--------|---------------|--------|----|----------|--------------------|
| 1          | 3989   | 100.0         | 4718   | 14 | AF063497 | AF063497 Adeno-ass |
| 2          | 3963   | 99.3          | 4683   | 14 | AF028704 | AF028704 Adeno-ass |
| 3          | 3511   | 88.0          | 4722   | 14 | AF028705 | AF028705 Adeno-ass |
| 4          | 3494   | 87.6          | 4726   | 14 | AVU48704 | U48704 Adeno-assoc |
| 5          | 3473   | 87.1          | 2211   | 14 | AY242997 | AY242997 Non-human |
| 6          | 3448   | 86.4          | 2217   | 14 | AY243015 | AY243015 Non-human |
| 7          | 3442.5 | 86.3          | 4721   | 14 | AF513851 | AF513851 Adeno-ass |
| 8          | 3439   | 86.2          | 8179   | 6  | AR264580 | AR264580 Sequence  |
| 9          | 3438   | 86.2          | 2217   | 14 | AY243007 | AY243007 Non-human |
| 10         | 3434   | 86.1          | 2187   | 14 | AY243006 | AY243006 Non-human |
| 11         | 3432   | 86.0          | 2217   | 14 | AY243009 | AY243009 Non-human |
| 12         | 3429   | 86.0          | 2217   | 14 | AY243023 | AY243023 Non-human |
| 13         | 3424.5 | 85.8          | 2214   | 14 | AY243020 | AY243020 Non-human |
| 14         | 3423.5 | 85.8          | 2208   | 14 | AY243021 | AY243021 Non-human |
| 15         | 3422.5 | 85.8          | 2190   | 14 | AY242998 | AY242998 Non-human |
| 16         | 3422   | 85.8          | 2217   | 14 | AY243022 | AY243022 Non-human |
| 17         | 3420.5 | 85.7          | 2202   | 14 | AY243004 | AY243004 Non-human |
| 18         | 3420   | 85.7          | 2187   | 14 | AY243013 | AY243013 Non-human |
| 19         | 3417.5 | 85.7          | 2190   | 14 | AY242999 | AY242999 Non-human |
| 20         | 3416.5 | 85.6          | 2202   | 14 | AY243012 | AY243012 Non-human |
| 21         | 3414.5 | 85.6          | 2190   | 14 | AY243000 | AY243000 Non-human |
| 22         | 3414.5 | 85.6          | 2196   | 14 | AY243010 | AY243010 Non-human |
| 23         | 3413.5 | 85.6          | 2208   | 14 | AY243008 | AY243008 Non-human |
| 24         | 3408.5 | 85.4          | 2196   | 14 | AY243011 | AY243011 Non-human |
| 25         | 3406   | 85.4          | 2187   | 14 | AY243016 | AY243016 Non-human |
| 26         | 3405   | 85.4          | 4393   | 14 | AF513852 | AF513852 Adeno-ass |
| 27         | 3402.5 | 85.3          | 4072   | 6  | AR222046 | AR222046 Sequence  |
| 28         | 3402.5 | 85.3          | 4072   | 6  | AX205074 | AX205074 Sequence  |
| 29         | 3402.5 | 85.3          | 4679   | 6  | AX282480 | AX282480 Sequence  |
| 30         | 3402.5 | 85.3          | 4679   | 14 | AF043303 | AF043303 Adeno-ass |
| 31         | 3402.5 | 85.3          | 7327   | 12 | AF369963 | AF369963 Cloning v |
| 32         | 3402.5 | 85.3          | 7557   | 6  | AR222045 | AR222045 Sequence  |
| 33         | 3402.5 | 85.3          | 7557   | 6  | AX205073 | AX205073 Sequence  |
| 34         | 3402.5 | 85.3          | 8698   | 6  | AR222044 | AR222044 Sequence  |
| 35         | 3402.5 | 85.3          | 8698   | 6  | AX205072 | AX205072 Sequence  |
| 36         | 3402   | 85.3          | 2187   | 14 | AY243017 | AY243017 Non-human |
| 37         | 3398.5 | 85.2          | 2202   | 14 | AY243014 | AY243014 Non-human |
| 38         | 3391   | 85.0          | 2187   | 14 | AY243018 | AY243018 Non-human |
| 39         | 3386   | 84.9          | 2187   | 14 | AY243019 | AY243019 Non-human |
| 40         | 3360.5 | 84.2          | 4680   | 6  | AR028767 | AR028767 Sequence  |
| 41         | 3360.5 | 84.2          | 4680   | 6  | I62303   | I62303 Sequence 1  |
| 42         | 3322.5 | 83.3          | 4675   | 6  | AX135805 | AX135805 Sequence  |
| 43         | 3322.5 | 83.3          | 4675   | 6  | AX286292 | AX286292 Sequence  |
| 44         | 3322.5 | 83.3          | 4675   | 6  | BD094552 | BD094552 Method of |
| 45         | 3322.5 | 83.3          | 4675   | 14 | AA2CG    | J01901 Adeno-assoc |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ALIGNMENTS

|   |   |               |      |        |                 |
|---|---|---------------|------|--------|-----------------|
| AF063497                                      | AF063497  | 4718 bp       | DNA  | linear | VRL 27-APR-1999 |
| LOCUS   | Adeno-associated virus 1, complete genome.  |               |      |        |                 |
| DEFINITION                                    | Adeno-associated virus 1, complete genome.  |               |      |        |                 |
| ACCESSION                                     | AF063497  |               |      |        |                 |
| VERSION                                       | AF063497.1  | GI:4689096    |      |        |                 |
| KEYWORDS                                      |   |               |      |        |                 |
| SOURCE  | Adeno-associated virus 1  |               |      |        |                 |
| ORGANISM                                      | Adeno-associated virus 1  |               |      |        |                 |
| REFERENCE                                     | Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.   |               |      |        |                 |
| AUTHORS                                       | 1 (bases 1 to 4718)<br>Xiao,W., Chirmule,N., Berta,S.C., McCullough,B., Gao,G. and<br>Wilson,J.M.   |               |      |        |                 |
| TITLE   | Gene therapy vectors based on adeno-associated virus type 1   |               |      |        |                 |
| JOURNAL                                       | J. Virol. 73 (5), 3994-4003 (1999)  |               |      |        |                 |
| MEDLINE                                       | 99214338  |               |      |        |                 |
| PUBMED  | 10196295  |               |      |        |                 |
| REFERENCE                                     | 2 (bases 1 to 4718)   |               |      |        |                 |
| AUTHORS                                       | Xiao,W. and Wilson,J.M.   |               |      |        |                 |
| TITLE   | Direct Submision  |               |      |        |                 |
| JOURNAL                                       | Submitted (05-MAY-1998) IHGT, University of Pennsylvania, 3601<br>Spruce Street, Philadelphia, PA 19104, USA  |               |      |        |                 |
| FEATURES                                      | Location/Qualifiers   |               |      |        |                 |
| source  | 1..4718<br>/organism="Adeno-associated virus 1"<br>/mol_type="genomic DNA"<br>/db_xref="taxon:85106"<br>/note="AAV1"<br>335..2206<br>/codon_start=1<br>/product="nonstructural protein"<br>/protein_id="AAD27758.1"<br>/db_xref="GI:4689098"  |               |      |        |                 |
| CDS   | /translation="MPGFYEIVIKVPSDDLDEHLPGISDSFVSVVAEKEWELPPDSMD<br>LNLIEQAPLTVAEKLQDFLVQWRRVSKAPEALFFVQFEKGESEYFHLHLIVETTVKVS<br>MVLGRFLSQIRDKLVQTIYRGIEPTLPNWPFAVTKTRNGAGGKNKVVDCEYIPNLLPK.<br>TQPELQWANTNMEEYISACLNLAEKRLVAQHLTHVSQTQONKENLNPNSDAPVRS<br>KTSARYMELVGLVDRGITSEKQWIOEDQASYISFNAASRSRQIKAAALDNAGKIMAL<br>TKSAPDVLGPAPPADIKTNRIYRIELNGEYEPAYAGSVFLGWAQKRGKRNITWLFQ<br>PATTGKTNIABAIHAHVPFYGCNVNWTNENFPFNDQVDMVWEEGKMTAKRVESAKA<br>ILGSKVRVDQCKSSAQIDPTPVIIVTSNTNMCVIDGNSTTFEHQQPLQDRMFKFEL<br>TRRLEHDFGKVTQKEVFEFFRWAQDHTEVAHEFYVRKGGANKRPAPDDADKSEPKRA<br>CPSVADPSTSDAEGAPVDFADRYQNKSRHAGMLQMLFPCKTCERMNQNFICFTHGT<br>RDCSECFPGVSESQPVVRKRTYRKLCAIHLHLGRAPEIACSACDLVNVLDLDDCVSEQ"<br>2223..4433<br>/codon_start=1<br>/product="capsid protein"<br>/protein_id="AAD27757.1"<br>/db_xref="GI:4689097"<br>/translation="MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQKQDDGRG<br>LVLPGYKYLGPFGNLDKGEPVNAADAALEHDKAYDQQLKAGDNPYLRYNHADAEPQE<br>RLQEDTSFGGNLGRAVFOAKKRVLEPLGVEEAKTAPGKKRPVEQSPQEPDSSSIGIG<br>KTGOQPAKRLNFGTGDSEVPDQPLGEPATPAAVGPTTMASSGGAPMADNNEGA<br>DGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNHLYKQISSASTGASNDNHFGY<br>STPWGYDFNRFHCFSPRDWORLINNWGFRPKRLNFKLFNQKEVTTNDGVTTIA<br>NNLTSTVQVSDSEYQLPYVLGSAHQCLPPFPADVFMIPOYGYLTNLNGSQAVGRSS<br>FYCLEYFPSQMLRTGNFTFSYTFEEVPFHSYAHQSQSLDRMLNPLIDQYLYLNRTO<br>NQSGSAQNKDLDFSRGSPAGMSVQPKNWLPGCYRQQRVSKTKTDNNNSNFTWTGASK<br>YNLNGRESIINPCTAMASHKDDKEDKFFPMGSGVMI FGKESAGASNTALDNVMTDEEII<br>KATNPVATEREGTVAVNFQSSSTDPTGDVHAMGALPGMVWQDRDYLQGPIMAKIPH<br>TDGHFHPSPLMGGFGLKNPFPQILIKNTVPANPPAEFSATKFASFITQYSTGQSVSE<br>IEWELQKENSQRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL" |               |      |        |                 |
| BASE COUNT                                    | 1121 a 1393 c 1273 g 931 t  |               |      |        |                 |
| ORIGIN  |   |               |      |        |                 |
| Alignment Scores:                             |   |               |      |        |                 |
| Pred. No.:                                    | 2,14e-240   | Length:       | 4718 |        |                 |
| Score:  | 3989.00   | Matches:      | 736  |        |                 |
| Percent Similarity:                           | 100.00%   | Conservative: | 0    |        |                 |
| Best Local Similarity:                        | 100.00%   | Mismatches:   | 0    |        |                 |
| Query Match:                                  | 100.00%   | Indels:       | 0    |        |                 |
| DB:   | 14  | Gaps:         | 0    |        |                 |
| US-09-807-802A-13 (1-736) x AF063497 (1-4718) |   |               |      |        |                 |

|    |      |   |      |
|----|------|---|------|
| QY | 1    | MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg    | 20   |
| Db | 2223 | ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTCGC    | 2282 |
| QY | 21   | GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp    | 40   |
| Db | 2283 | GAGTGGTGGGACTTGAAACCTGGAGCCCCGAAAGCCCAAGCCAAAGCAGGAC            | 2342 |
| QY | 41   | AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp    | 60   |
| Db | 2343 | GACGGCCGGGGTCTGGTCTCTGGCTTACAGTACTCGGACCCCTTCAACGGACTCGAC       | 2402 |
| QY | 61   | LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp       | 80   |
| Db | 2403 | AAGGGGAGCCCGTCAACGGCGGACGACGCGGCCCTCGAGCAGCAAGGCGCTACGAC        | 2462 |
| QY | 81   | GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe    | 100  |
| Db | 2463 | CAGCAGCTCAAAGCGGTGACAATCCGTACCTGCGGTATAACCAACGCGCGCGAGTTT       | 2522 |
| QY | 101  | GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln    | 120  |
| Db | 2523 | CAGGAGCGTCTGCAAGAAGATACGTCTTTTGGGGCAACCTCGGGCGAGCAGTCTCCAG      | 2582 |
| QY | 121  | AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro       | 140  |
| Db | 2583 | GCCAAGAACGGGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGCTCCT     | 2642 |
| QY | 141  | GlyLysLysArgProValGlnSerProGlnGluProAspSerSerSerGlyIleGly       | 160  |
| Db | 2643 | GGAAAGAAACGTCGCGGTAGAGCAGTCGCCACNAGAGCCAGACTCCTCCTCGGGCATCGGC   | 2702 |
| QY | 161  | LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu    | 180  |
| Db | 2703 | AAGACAGGCCAGCAGCCCCGCTAAAAAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG   | 2762 |
| QY | 181  | SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaValGlyPro       | 200  |
| Db | 2763 | TCAGTCCCCGATCCACAACCTCTCGGAGAACCTCCAGCAACCCCGCTGCTGTGGGACCT     | 2822 |
| QY | 201  | ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly       | 220  |
| Db | 2823 | ACTACAATGGCTTCAGGGCGTGGCGCACCAATGGCAGACAATAACGAAGCGCCGACGGA     | 2882 |
| QY | 221  | ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle    | 240  |
| Db | 2883 | GTGGGTAAATGCCTCAGGAAATTTGGCATTTCCGATTTCCACATGGTGGGCGACAGAGTCATC | 2942 |
| QY | 241  | ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle    | 260  |
| Db | 2943 | ACCACCAGCACCCGACCTGGGCTTGCCCACTTACAATAACCACTCTACAAGCAAATC       | 3002 |
| QY | 261  | SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp    | 280  |
| Db | 3003 | TCCAGTGTCTCAACGGGGGCCAGCAACGACAACCACTACTTCGGCTACAGCACCCCTGG     | 3062 |
| QY | 281  | GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu    | 300  |
| Db | 3063 | GGGTATTTTGAATTCACACAGATTCCACTGCCACTTTTCCACCACGTGACTGGCAGCGACTC  | 3122 |
| QY | 301  | IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln    | 320  |
| Db | 3123 | ATCAACAACAATTTGGGATTTCCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA  | 3182 |
| QY | 321  | ValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThr    | 340  |
| Db | 3183 | GTCAAGGAGGTACGACGAATGATGCGCTCACAACTATCGCTAATAACCTTACCAGCAGC     | 3242 |
| QY | 341  | ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln    | 360  |
| Db | 3243 | GTTCAAGTCTTCTCGGACTCGGAGTACCAGCTTCCGTACGTCTCGGCTCTGGCACCAG      | 3302 |



Qy 361 GlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr 380  
Db 3303 GGCTGCCTCCCTCCGTTCCCGCGGACGTTTCATGATTCGCAATACGGCTACCTGACG 3362  
Qy 381 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 400  
Db 3363 CTCACAATGGCAGCCAGCCGTTGGGACGTTTCATCCTTTTACTGCCTGGAATATTTCCCT 3422  
Qy 401 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro 420  
Db 3423 TCTCAGATGCTGAGAACGGCAACAACCTTTTACTCTCAGCTACACCTTTGAGGAAGTGCCT 3482  
Qy 421 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 440  
Db 3483 TTCCACAGCAGCTACCGGCACAGCCAGCCTGGACCGGCTGATGAATCCTCTCATCGAC 3542  
Qy 441 GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAsp 460  
Db 3543 CAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAACAACAGGAC 3602  
Qy 461 LeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuPro 480  
Db 3603 TTGCTGTTTAGCCGTGGTCTCCAGCTGGCATGTCTGTTTCAGCCCCAAAACACTGGCTACCT 3662  
Qy 481 GlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsn 500  
Db 3663 GGACCCTGTTATCGGCAGCAGCGCGTTCCTTCTAAAAACAAAACAGACAAACAACAGCAAT 3722  
Qy 501 PheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnPro 520  
Db 3723 TTTACCTGGACTGGTGCTTCAAAATATAACCTCAATGGCGTGATCCATCATCAACCTT 3782  
Qy 521 GlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyVal 540  
Db 3783 GGCACCTGCTATGGCCTCACACAAGACGACGAGCAAGATTCTTTCCCATGACGGTGTC 3842  
Qy 541 MetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIle 560  
Db 3843 ATGATTTTGGAAAAGAGAGCGCGGAGCTTCAAACTGCTGATGGACAATGTCATGATT 3902  
Qy 561 ThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrVal 580  
Db 3903 ACAGACGAAGAGGAATTAAGCCACTAACCTCTGGCCACCGAAAGATTGGGACCGTG 3962  
Qy 581 AlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGly 600  
Db 3963 GCAGTCAATTTCCAGAGCAGCAGCACAGACCTCGGACCGGAGATGTGCATGTATGGGA 4022  
Qy 601 AlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAla 620  
Db 4023 GCATTACCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCAGGGTCCCATTTGGCC 4082  
Qy 621 LysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeu 640  
Db 4083 ANAATTCCTCACACAGATGGACACTTTTACCCCGTCTCCTCTTATGGCGGCTTTGGACTC 4142  
Qy 641 LysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAla 660  
Db 4143 AAGAACCCGCCTCCTCAGATCCTCATCAAAAACACGCTGTCTCTGCGAATCCTCCGCG 4202  
Qy 661 GluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSer 680  
Db 4203 GAGTTTTCAGCTACAAAGTTTGCTTTCATTCATCAACCAATACTCCACAGACAAGTGAGT 4262  
Qy 681 ValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGln 700  
Db 4263 GTGGAAATTGAATGGGAGCTGCAGAAAGAAAACAGAACGCTGGAATCCCGAAGTGCAG 4322  
Qy 701 TyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeu 720  
Db 4323 TACACATCCAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACACAATGGACTT 4382  
Qy 721 TyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736

Db 4383 TATACTGAGCCTCGCCCCATTGGCACCCGTTACCTTACCTACCGTCCCCTG 4430  
RESULT 2  
AF028704  
LOCUS AF028704 4683 bp DNA linear VRL 12-JAN-1998  
DEFINITION Adeno-associated virus 6, complete genome.  
ACCESSION AF028704  
VERSION AF028704.1 GI:2766605  
KEYWORDS  
SOURCE  
ORGANISM Adeno-associated virus 6  
Adeno-associated virus 6  
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.  
REFERENCE 1 (bases 1 to 4683)  
AUTHORS Rutledge,E.A., Halbert,C.L. and Russell,D.W.  
TITLE Infectious clones and vectors derived from adeno-associated virus (AAV) serotypes other than AAV type 2  
JOURNAL J. Virol. 72 (1), 309-319 (1998)  
MEDLINE 98080418  
PUBMED 9420229  
REFERENCE 2 (bases 1 to 4683)  
AUTHORS Rutledge,E.A. and Russell,D.W.  
TITLE Direct Submission  
JOURNAL Submitted (24-SEP-1997) Hematology, Univ. of Washington, Box 357720, Seattle, WA 98195, USA  
FEATURES  
source  
1. .4683  
/organism="Adeno-associated virus 6"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:68558"  
/note="AAV6"  
320. .2191  
/codon\_start=1  
/evidence=not\_experimental  
/product="nonstructural protein Rep78"  
/protein\_id="AAB95449.1"  
/db\_xref="GI:2766606"  
/translation="MPGFYEIVIKVPSDLDEHLPGLSDSFVNWVAEKEWELPPDSMD  
LNLIEQAPLTVAEKLQDFLVQRRVSKAPEALFVQPEKGESYFHLHLVETTVGKS  
MVLGRFLSQIRDKLVQTIYRGIEPTLPNWFVAVTKRNGAGGKVKVDECYIPNYLLPK  
TOPELQAWTNMEEYISACNLAEKRLVAHDLTHVSTQEOQNKENLNPNSDAPVRS  
KTSARYMELVGLVDRGITSEKQWQEDQASYSFNAASNSRSQIKAALDNAGKTMAL  
TKSAPDYLVGPPAPPADIKTNRIYRILELNGYDPAVAGSVPLGWAQKRFGRNTIWLFG  
PATGKTNIAEIAHAIVPFYGCNVNTNENFPFNDVCDVKWVIWEEGKMTAKVVESAKA  
ILGSKVRVDQCKSSAQIDPTPVIVTSNTNMCAVDGNSTTFEHQQPLQDRMPKPEL  
TRRLEHDFGKVTKEVKEFFRWAQDHVTEVAHEFVYRKGANKRPAPDDADKSEPKRA  
CPSVADPSTDAEGAPVDADRYQNKCSRHAGLMQLFPCKTCERMNQNFNCFTHGT  
RDCSECFPGVSESQPVVRKRTYRKLCAIHLLGRAPEIACSDLVNVDLDDCVSEQ"  
2208. .4418  
/codon\_start=1  
/evidence=not\_experimental  
/product="capsid protein VP1"  
/protein\_id="AAB95450.1"  
/db\_xref="GI:2766607"  
/translation="MAADGYLPDWLEDNLSEGIREWWDLPKGA PKPAKQKQDDGRG  
LVLPGYKYLGPFGNLDKGEVPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEFQE  
RLQEDTSFGNLRGAVQAQKRVLEPFGLVBEAGATAPGKXRPVEQSPQEPDSSSGIG  
KTGQPAKKRLNFGTGDSVDPDQPLGEPPTAAVGPPTMASGGGAPMADNNEGA  
DGVGNASGNWHCDSTWGLDRVITTSRTWALPTVNNHLYKQISSASTGASNDNHYFGY  
STPWGYFDENRFCHFSPRDWQRLINNNGWFRPKLNFKLFIQVKEVTTNDGVTTIA  
NNLTSTVQVFSQSEYQLPYVLGSAHQCLPPFPADVPMI PQYGYLTNLNGSQAVGRSS  
FYCLEYFPPSQMLRTGNFTPSYTFEDVPFHSSYAHQSRLDMLNPLIDQYLYLNRQT  
NQSGAQNKKOLLFSRGS PAGMSVQPKNWLPGCPYQVRVSKRTKTDNNNSNFTWTGASK  
YNLNGRESIIINPGTAMASHKDDKDKFFPMSGVMIFGKESAGASNTALDNVMTDEEI  
KATNPVATERFGTIVNLQSSSTD PATGDVHVMGALPGMVWQDRDVYLOGPIWAKIPH  
TDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPAEFSATKFPASFITQYSTGQSVSE  
IEWELQKENSKRWNEVQYTSNYAKSANVDFTVDDNNGLYTEPRPIGTRYLTRPL"

BASE COUNT 1114 a 1363 c 1277 g 929 t

ORIGIN

Alignment Scores:

Pred. No.: 9.01e-239 Length: 4683

Score: 3963.00 Matches: 730

|   |      |  |      |                 |  |  |  |
|---|------|--|------|-----------------|--|--|--|
| Percent Similarity: 99.59%                    |      |  |      | Conservative: 3 |  |  |  |
| Best Local Similarity: 99.18%                 |      |  |      | Mismatches: 3   |  |  |  |
| Query Match: 99.35%                           |      |  |      | Indels: 0       |  |  |  |
| DB: 14  |      |  |      | Gaps: 0         |  |  |  |
| US-09-807-802A-13 (1-736) x AF028704 (1-4683) |      |  |      |                 |  |  |  |
| QY  | 1    | MetAlaAlaAspGlyTyrLeuProAspTyrLeuGluAspAsnLeuSerGluGlyIleArg   | 20   |                 |  |  |  |
| Db  | 2208 | ATGSGTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC    | 2267 |                 |  |  |  |
| QY  | 21   | GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp   | 40   |                 |  |  |  |
| Db  | 2268 | GAGTGGTGGACTTGAAACCTGGAGCCCCGAAACCCAAAGCCAAACAGCAGGAC          | 2327 |                 |  |  |  |
| QY  | 41   | AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp   | 60   |                 |  |  |  |
| Db  | 2328 | GACGGCCGGGTCTGGTCTTCTCGGTACAAGTACCTCGGACCTTCAACGGGACTCGAC      | 2387 |                 |  |  |  |
| QY  | 61   | LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp      | 80   |                 |  |  |  |
| Db  | 2388 | AAGGGGAGCCCGTCAACCGCGGGATGCGAGCCCTCGAGCACGACACAGGCTACGAC       | 2447 |                 |  |  |  |
| QY  | 81   | GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe   | 100  |                 |  |  |  |
| Db  | 2448 | CAGCAGCTCAAAGCGGTGACAATCCGTACCTGCGGTATAACCAACGCCGACGCCGAGTTT   | 2507 |                 |  |  |  |
| QY  | 101  | GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln   | 120  |                 |  |  |  |
| Db  | 2508 | CAGGAGCGTCTGCAAGAAGATACGTCTTTTGGGGCAACCTCGGGCGAGCAGTCTTCCAG    | 2567 |                 |  |  |  |
| QY  | 121  | AlaLysLysArgValLeuGluProLeuGlyLeuValGluGluGlyAlaLysThrAlaPro   | 140  |                 |  |  |  |
| Db  | 2568 | GCCAAAGAAGAGGGTTCTCGAACCTTTTGGTCTGGTTGAGGAAGTGTCTAAGACGGCTCCT  | 2627 |                 |  |  |  |
| QY  | 141  | GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly   | 160  |                 |  |  |  |
| Db  | 2628 | GGAAGAAACGTCCGGTAGAGCAGTCGCCACAGAGCCAGACTCCTCTCTCGGGCATTTGC    | 2687 |                 |  |  |  |
| QY  | 161  | LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu   | 180  |                 |  |  |  |
| Db  | 2688 | AAGACAGGCCAGCAGCCCGCTAAAAAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG   | 2747 |                 |  |  |  |
| QY  | 181  | SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro   | 200  |                 |  |  |  |
| Db  | 2748 | TCAGTCCCCGACCCACAACCTCTCGGAGAACCTCCAGCAACCCCGCTGCTGTGGGACCT    | 2807 |                 |  |  |  |
| QY  | 201  | ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly      | 220  |                 |  |  |  |
| Db  | 2808 | ACTACAATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAAATAACGAAGGCCCGCAGCGA  | 2867 |                 |  |  |  |
| QY  | 221  | ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle   | 240  |                 |  |  |  |
| Db  | 2868 | GTGGGTAATGCCCTCAGGAAATTGGCATTTGCCATTCCACATGGCTGGGCGACAGAGTCATC | 2927 |                 |  |  |  |
| QY  | 241  | ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle   | 260  |                 |  |  |  |
| Db  | 2928 | ACCACGACACCCGAAACATGGGCCTTGCCACCTATAACAACCCACCTCTACAAGCAAAATC  | 2987 |                 |  |  |  |
| QY  | 261  | SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp   | 280  |                 |  |  |  |
| Db  | 2988 | TCCAGTGTCTTCAACGGGGGCCAGCAACGACCAACCCTACTTCCGCTACAGCACCCCTTGG  | 3047 |                 |  |  |  |
| QY  | 281  | GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu   | 300  |                 |  |  |  |
| Db  | 3048 | GGGTATTTTGATTTTCAACAGATTCCACTGCCATTCTCACCACGTGACTGGCAGCGACTC   | 3107 |                 |  |  |  |
| QY  | 301  | IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln   | 320  |                 |  |  |  |
| Db  | 3108 | ATCAACAACAATTGGGGATTCCGGGCCCAAGAGACTCAACTTCAAGCTCTTCAACATCCAA  | 3167 |                 |  |  |  |
| QY  | 321  | ValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThr   | 340  |                 |  |  |  |

|    |      |  |      |  |  |  |  |
|----|------|--|------|--|--|--|--|
| Db | 3168 | GTCAAGGAGGTCAAGACGAATGATGGCGTCACGACCATCGCTAATAACCTTACCAGCAGC | 3227 |  |  |  |  |
| QY | 341  | ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln | 360  |  |  |  |  |
| Db | 3228 | GTTCAAGTCTTCTCGGACTCGGAGTACCAGTTGCCGTACGTCTCGGCTCTGCGCACCAG  | 3287 |  |  |  |  |
| QY | 361  | GlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr    | 380  |  |  |  |  |
| Db | 3288 | GGCTGCTCCCTCCGTTCCCGCGGACGTGTTTATGATTCGCGAGTACGGCTACCTAACG   | 3347 |  |  |  |  |
| QY | 381  | LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro | 400  |  |  |  |  |
| Db | 3348 | CTCAACAATGGCAGCCAGGCGAGTGGGACGGTCACTCTTTTACTGCTGGAATATTTCCCA | 3407 |  |  |  |  |
| QY | 401  | SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro | 420  |  |  |  |  |
| Db | 3408 | TCGCAGATGCTGAGAACCGGCAATAACTTTACCTTCAGCTACACCTTCGAGGACGTGCCT | 3467 |  |  |  |  |
| QY | 421  | PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp | 440  |  |  |  |  |
| Db | 3468 | TTCCACAGCAGCTACCGGCACAGCCAGAGCCTGGACCGGTGATGAATCTCTCATTCGAC  | 3527 |  |  |  |  |
| QY | 441  | GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAsp | 460  |  |  |  |  |
| Db | 3528 | CAGTACCTGTATTACCTGAACAGAACTCAGAAATCAGTCCGGAAGTCCCCAAACAAGGAC | 3587 |  |  |  |  |
| QY | 461  | LeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuPro | 480  |  |  |  |  |
| Db | 3588 | TTGCTGTTAGCCGGGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAAACCTGGTACCT   | 3647 |  |  |  |  |
| QY | 481  | GlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnSerAsn    | 500  |  |  |  |  |
| Db | 3648 | GGACCTGTATTACCGGCAGCAGCGCGTTCTTAAACAAAAACAGACAACAACAGCAAC    | 3707 |  |  |  |  |
| QY | 501  | PheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnPro | 520  |  |  |  |  |
| Db | 3708 | TTTACCTGGACTGGTGTCTTCAAAATATAACCTTAATGGCGTGGAATCTATAATCAACCT | 3767 |  |  |  |  |
| QY | 521  | GlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyVal | 540  |  |  |  |  |
| Db | 3768 | GGCACTGTATGGCCTCACACAAGACGACAAAGACAAGTTCTTCCCATGAGCGGTGTC    | 3827 |  |  |  |  |
| QY | 541  | MetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIle | 560  |  |  |  |  |
| Db | 3828 | ATGATTTTGGAAAGGAGAGCGCCGAGCTTCAAACACTGCATTTGGACAATGTCTATGATC | 3887 |  |  |  |  |
| QY | 561  | ThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrVal | 580  |  |  |  |  |
| Db | 3888 | ACAGACGAAGAGGAAATCAAAGCCACTAACCCCGTGGCCACCGAAAGATTTGGGACTGTG | 3947 |  |  |  |  |
| QY | 581  | AlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGly | 600  |  |  |  |  |
| Db | 3948 | GCAGTCAATCTCCAGAGCAGCAGCAGACACCTCGACCCGAGATGTGCATGTTATGGGA   | 4007 |  |  |  |  |
| QY | 601  | AlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAla | 620  |  |  |  |  |
| Db | 4008 | GCCTTACCTGGAATGGTGTGGCAAGACAGAGACGTATACCTGCAGGTCCTATTTGGGCC  | 4067 |  |  |  |  |
| QY | 621  | LysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyPheGlyLeu    | 640  |  |  |  |  |
| Db | 4068 | AAAATTCTCACACGGATGGACACTTTCACCCGTCTCTCTCATGGCGGCTTTGGACTT    | 4127 |  |  |  |  |
| QY | 641  | LysAsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAla | 660  |  |  |  |  |
| Db | 4128 | AAGCACCCGCTCTCAGATCCTCATCAAAAACACGCCCTGTTCTCTGCGAATCTCTCCGCA | 4187 |  |  |  |  |
| QY | 661  | GluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSer | 680  |  |  |  |  |
| Db | 4188 | GAGTTTTCGGCTACAAAAGTTTGTCTTCATTCATCACCAGTATTTCCAGGACAAGTGAGC | 4247 |  |  |  |  |
| QY | 681  | ValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGln | 700  |  |  |  |  |
| Db | 4248 | GTGGAGATTGATGGGAGCTGCAGAAAGAAAAACAGCAACCGCTGGAATCCCGAAGTGACG | 4307 |  |  |  |  |

```
QY 701 TyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeu 720
Db 4308 TATACATCTAACTATGCAAAATCTGCCAACGTTGATTCTCACTGTGGACAACAATGGACTT 4367

QY 721 TyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736
Db 4368 TATACTGAGCCTCGCCCATTTGGCACCCGTTACCTCACCCGTCCTCCCTG 4415

RESULT 3
AF028705 AF028705 4722 bp DNA linear VRL 12-JAN-1998
LOCUS Adeno-associated virus 3B, complete genome.
DEFINITION AF028705
ACCESSION AF028705
VERSION AF028705.1 GI:2766608
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 4722)
AUTHORS Rutledge,E.A., Halbert,C.L. and Russell,D.W.
TITLE Infectious clones and vectors derived from adeno-associated virus
(AAV) serotypes other than AAV type 2
JOURNAL J. Virol. 72 (1), 309-319 (1998)
MEDLINE 98080418
PUBMED 9420229
REFERENCE 2 (bases 1 to 4722)
AUTHORS Rutledge,E.A. and Russell,D.W.
TITLE Direct Submission
JOURNAL Submitted (25-SEP-1997) Hematology, Univ of Washington, Box 357720,
Seattle, WA 98195, USA

FEATURES
source
1. .4722
/organism="Adeno-associated virus 3B"
/mol_type="genomic DNA"
/db_xref="taxon:68742"
/note="AAV3B"
317. .2191
/codon_start=1
/product="nonstructural protein Rep78"
/protein_id="AAB95451.1"
/db_xref="GI:2766609"
/translation="MPGFYEIVLKVPSLDLDEHLPGISNFSFVNWVAEKWELPPDSMD
PNLIEQAPLTVAEKLQREFLVEWRVRSKAPEALFFVQFEKGETYFHLHLVLIETIGVKS
MVVGRYVSQIKEKLVTIRYRGVEPQLPNWFVATKTRNGAGGKNVDDCYIPNYLLPK
TOPELOWANTMDQYLSACNLAEKRILVAQHLTHVSQTQEQNKQNPNSDAPVIRS
KTSARYMELVGLVDRGITSEKQWIEDQASYISFNAASNSRSQIKAALDNASKIMSL
TKTAPDYLGSNPPEDITKNRIYQILELNGYDPQYAAVSFLGWAQKFKGRNTIWLFG
PATTGKTNIABAIHAHAFYFGCVNWTNENFPNDCVDMKVIWEECKMTAKVVESAKA
ILGSKVRVDQCKSSAQIEPTPIVTSNTNCAVIDGNSITFEHQQLQDRMEKFEF
TRRLDHDGKVTQKQEVKOFFRWASDHVTDVAHEFYVRKGAKRPASNDADVSEPKRQ
CTSLAQPTTSDAEAPADYADRYQNKCSRHVGMNLMFLPCKTCERMNQISNVCFTHGQR
DCGECFPGMSQSPSVVKKTYQKLCPIHILGRAPEIACSACDLANVDLDDCVSEQ
"
2208. .4418
/codon_start=1
/product="capsid protein VP1"
/protein_id="AAB95452.1"
/db_xref="GI:2766610"
/translation="MAADGYLPDWLEDNLSEGIREWALKPGVPQPKANQQHQDNRRG
LVLPGYKYLPGNGLDKGEPVNEADAALEHDKAYDQQLKAGDNPYLKYNHADAEEFQE
RLQEDTSFGNIGRAVFQAKRIILEPLGLVEEAAKTAPKKRPVDPQSPQEPDSSGVG
KSGKQAPARKLNFQGTGDSVDPQPLGEPFPAAPTSLGNTWASGGGAPMADNNEGA
DVGNSSGNWHCDSQWLGDRIITSTRTWALPTYNNHLYKQISSQSGASNDNHIFYGS
TPWGYDFDNRFCHFSPRDWORLINNWGFRPKLSFKLENIQVKEVTQNDGTTTIAN
NLSTVQVFTDSEYQLPYVLGSAHQGLPPFPADVFMVPOYGYLTNNGSQAVGRSSF
YCLEYFPPSOMLRTGNFQFSYTFEDVPFHSSYAHQSQSLDRMLNPLIDQYLYLNRTOG
TTSGTTNQSRLLFSQAGPQMSLQARNWLPQCYRQQRLSKTANDNNNSNFPWTAASK
YHLNGRDLSVNPFPAMASHKODEKFFPMHGNLI FGKEGTASNALDNVMITDEEEI
RTTNPVATEQYGTVANNLQSSNTAPTTRTVNDQGLPGMWQDRDVLQGPWIKIPH
TDGHFHPSPLMGGFGLKHPPPQIMIKNTVPVNPANPTTFSPAKFASFTQYSTQVSVSE
IEWELQKENSKRWNPEIQYTSNYSKSVNVDFTVDTNGVYSEPRPIGTRYLTRNL"
BASE COUNT 1204 a 1297 c 1243 g 978 t
```

ORIGIN

```
Alignment Scores:
Pred. No.: 1.88e-210 Length: 4722
Score: 3511.00 Matches: 640
Percent Similarity: 91.99% Conservative: 38
Best local Similarity: 86.84% Mismatches: 57
Query Match: 88.02% Indels: 2
DB: 14 Gaps: 2

US-09-807-802A-13 (1-736) x AF028705 (1-4722)

QY 1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20
Db 2208 ATGGCTGCTGACGGTTATCTTCCAGATTGGCTCGAGGACAACCTTTCTGAAGGCATTCTG 2267

QY 21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnLysGlnAsp 40
Db 2268 GAGTGGTGGGCTCTGAACCTGGAGTCCCTCAACCCAAAGCGAACCACCAACACCAGGAC 2327

QY 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp 60
Db 2328 AACCGTCGGGGTCTTTGTGCTTCCGGGTACAAATACCTCGGACCCGGTAACGGACTCGAC 2387

QY 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80
Db 2388 AAAGGAGAGCCGGTCAACGAGGCGGACGCGGAGCCCTCGAAACACGACAAAGCTTACGAC 2447

QY 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100
Db 2448 CAGCAGCTCAAGGCCGGTGACAAACCCGTACCTCAAGTACAACCAACGCGCGGAGTTT 2507

QY 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120
Db 2508 CAGGAGCGTCTTCAAGAAGATACGTCTTTTGGGGCAACCTTGGCAGAGCAGTCTTCCAG 2567

QY 121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro 140
Db 2568 GCCAAAAGAGGATCCTTGAGCCTCTTGGTCTGGTTGAGGAAGCAGCTAAAACGGCTCCT 2627

QY 141 GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly 160
Db 2628 GGAAAGAAGAGGCCCTGTAGATCAGTCTCTCTCAGGAACCGGACTCATCATCTTGGTGTGGC 2687

QY 161 LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180
Db 2688 AAATCGGGCAACAGCCCTGCCAGAAAAGACTAAATTTGGGTCAAGTCTGGCGACTCAGAG 2747

QY 181 SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro 200
Db 2748 TCAGTCCCAGACCCTCAACCTCTCGGAGAACCCACGAGCCCCCAAGTTTGGGATCT 2807

QY 201 ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly 220
Db 2808 AATACAATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAGGGTGC CGATGGA 2867

QY 221 ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle 240
Db 2868 GTGGGTAATTCTCTCAGGAAATTGGCATTGGGATTCCCAATGGCTGGCGCAGAGTCATC 2927

QY 241 ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle 260
Db 2928 ACCACCAGCACCAAGAACCTGGGGCCCTGCCACTTACAACAACCATCTCTACAAGCAATC 2987

QY 261 SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp 280
Db 2988 TCCAGCCAATCA--GGAGCTTCAAACGACAACCACTACTTTGGCTACAGACCCCTTGG 3044

QY 281 GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu 300
Db 3045 GGGTATTTTGACTTTAAACAGATTCCACTGCCACTTCTCACCACGCTGACTGGCAGCGACTC 3104

QY 301 IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln 320
```



Db 3105 ATTAACAACAACACTGGGGATTCCGGCCCAAGAACTCAGCTTCAAGCTCTTCAACATCCAA 3164  
Qy 321 VallysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThr 340  
Db 3165 GTTAAAGAGGTCACGCAGAACGATGGCAGCAGACTATTGCCAATAACCTTACCAGCAGC 3224  
Qy 341 ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln 360  
Db 3225 GTTCAAGTGTTTACGGACTCGGAGTATCAGCTCCCGTACGTGCTCGGGTCGGCGCACCA 3284  
Qy 361 GlyCysLeuProPropheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr 380  
Db 3285 GGCTGTCTCCCGCGTTTCCAGCGGACGTCTTTCATGTGTCCTCAGTATGGATACCTCACC 3344  
Qy 381 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 400  
Db 3345 CTGAACAACGGAAGTCAAGCGGTGGACGCTCATCTTTTACTGCTGAGTACTTCCCT 3404  
Qy 401 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro 420  
Db 3405 TCGCAGATGCTAAGGACTGGAATAACTTCCAATTTCAGCTATACCTTCGAGGATGACCT 3464  
Qy 421 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 440  
Db 3465 TTTTCACAGCAGCTACGCTCACAGCCAGAGTTTGGATCGCTTGATGAATCCTCTTATTGAT 3524  
Qy 441 GlnTyrLeuTyrTyrLeuAsnArgThrGln---AsnGlnSerGlySerAlaGlnAsnLys 459  
Db 3525 CAGTATCTGTACTACTCTGAACAGACGCAAGGAACAACCTCTGGAACAACCAACCAATCA 3584  
Qy 460 AspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeu 479  
Db 3585 CGGCTGCTTTTAGCCAGGCTGGGCTCAGTCTATGTCTTTCAGCGCCAGAAATTGGCTA 3644  
Qy 480 ProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSer 499  
Db 3645 CCTGGGCCCTGCTACCGGCAACAGAGACTTTCAAAGACTGTCAACGACAACAACAACAGT 3704  
Qy 500 AsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsn 519  
Db 3705 AACTTTCTTGGACAGCGGCCAGCAAAATATCATCTCAATGGCCGCGACTCGCTGGTGAAT 3764  
Qy 520 ProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGly 539  
Db 3765 CCAGGACCAGCTATGGCCAGTCAACAGGACGATGAAGAAATAATTTTCCCTATGCACGGC 3824  
Qy 540 ValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMet 559  
Db 3825 AATCTAATATTGGCAAGAAGGGAACAACGGCAAGTAACCGAGAATTAGATAATGTAATG 3884  
Qy 560 IleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThr 579  
Db 3885 ATTACGGATGAAGAAGAGATTTCGTACCACCAATCCTGTGGCAACAGACAGTATGGAAC 3944  
Qy 580 ValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMet 599  
Db 3945 GTGGCAATAACTTGCAGAGCTCAAATACAGCTCCCACGACTAGAAGTGTCAATGATCAG 4004  
Qy 600 GlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrp 619  
Db 4005 GGGGCCTTACCTGGCATGGTGTGGCAAGATCGTGACGTGTACCTTCAAGGACCTATCTGG 4064  
Qy 620 AlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGly 639  
Db 4065 GCAAAGATTCTCACACGGATGGACACTTTTCATCTCTCCTCTGATGGGAGGCTTTGGA 4124  
Qy 640 LeuLysAsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProPro 659  
Db 4125 CTGAACATCCGCCCTCCTCAAATCATGATCAAAAATACTCCGGTACCGCAAAATCCTCCG 4184  
Qy 660 AlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnVal 679

Db 4185 ACGACTTTTCAGCCCGGCCAAGTTTGTCTTCATTTATCACTCAGTACTCCACTGGACAGGTC 4244  
Qy 680 SerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluVal 699  
Db 4245 AGCGTGGAATTTGAGTGGGAGCTACAGAAAGAAAACAGCAAAACGTTTGGAAATCCAGAGATT 4304  
Qy 700 GlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGly 719  
Db 4305 CAGTACACTTCCAACCTACAACAAGTCTGTAAATGTGGACTTTTACTGTAGACACTAATGGT 4364  
Qy 720 LeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736  
Db 4365 GTTTATAGTGAACCTCGCCCTATTGGAACCCGCTATCTCACACGAAACTTG 4415  
RESULT 4  
AVU48704 4726 bp DNA linear VRL 15-JUL-1996  
LOCUS Adeno-associated virus 3 nonstructural protein and capsid protein  
DEFINITION genes, complete cds, and complete genome.  
ACCESSION U48704  
VERSION U48704.1 GI:1408467  
KEYWORDS Adeno-associated virus 3  
SOURCE Adeno-associated virus 3  
ORGANISM Adeno-associated virus 3  
REFERENCE 1 (bases 1 to 4726) Muramatsu, S., Mizukami, H., Young, N.S. and Brown, K.E.  
AUTHORS Nucleotide sequencing and generation of an infectious clone of  
TITLE adeno-associated virus 3  
JOURNAL Virology 221 (1), 208-217 (1996)  
MEDLINE 96266430  
PUBMED 8661429  
REFERENCE 2 (bases 1 to 4726) Muramatsu, S. and Brown, K.E.  
AUTHORS Direct Submission  
TITLE Submitted (06-FEB-1996) Shin-ichi Muramatsu, Hematology Branch,  
JOURNAL NHLBI/NIH, 9000 Rockville, MD 20892, USA  
FEATURES Location/Qualifiers  
1. .4726  
/organism="Adeno-associated virus 3"  
/mol\_type="genomic DNA"  
/strain="3H"  
/db\_xref="taxon:46350"  
318. .2192  
/codon\_start=1  
/product="nonstructural protein"  
/protein\_id="AAC55048.1"  
/db\_xref="GI:1408468"  
/translation="MPGFYEIVLKVPISDLERLPGISN FVNWVAEKWDVPPDSMD  
PNLIEQAPLTVAEKLQREFLVEWRVSKAPEALFFVQFEKGETYFHLVLIETIGVKS  
MVGVRYVSIKEKLVTRIYRGVEPOLPNWFVAVTKTRNGAGGKNKVDDCYIPNYLLPK  
TQPELQAWTNMDQYLSACLNLAEKRLVAQHLTHVSQTQEQNKENQNPNSDAPVIRS  
KTSARYMELVGLVDRGITSEKQWIQEDQASYISFNAASNSRSQIKAAALDNASKIMSL  
TKTAPDYLVSNNPPEDITKNRIYQIIEELNGYDPQYAA SVFLGWAQKKFGKRNITWLFG  
PATTKTNIAEAI AHAVPFYGCNVNTNENFPFND CVDKMWIWEEGKMTAKVESAKA  
ILGGSKVRVDQKCKSSAQIEPTPVITSNMCAVIDGNSSTTPEHOQPLQDRMFBFEL  
TRRLDHPGKVTQEVKQDFFRWASDHVTDVAHEFYVRKGGAKKRPASNDADVSEPKRE  
CTSLAQPTTSDAEAPADYADRYONKCSRHVGMNLM LFPCKTCERNQIISNVCFTHGQR  
DCGECFPGMSESQPVSVVKKTYQKLCPIHHILGRAPEIACSACDLANVDLDDCVSEQ  
"  
2209. .4419  
/codon\_start=1  
/product="capsid protein"  
/protein\_id="AAC55049.1"  
/db\_xref="GI:1408469"  
/translation="MAADGYLPDWLEDNLSEGI REWWALKPGVPQPKANQQHQDNRRR  
LVLPGYKILGPGNGLDKGEPVNEADAALEHDKAYDQQLKAGDNPYLKNHADAEFQE  
RLQEDTSFGNGLGRAVFAQKRILEPLGLVEEAAKTA PGKKGAVDQSPQEPDSSSGVG  
KSGKQAPARKRLNFGQTDSESVDPQPLGEP PPAAPTSLGNSNTWASGGAPMADNNEGA  
DGVNSSGNWHCDQSQWLGDRIYITSTRT WALPTYNNHLYKQISSGASNDNHYFGYS  
TPWGYFDNRFCHFSPRDQWRLINNNWGF RPKKLSFKLEINIQVRGVTQNDGTTIAN  
NLSTSTVQFTDSEYQLPYVLGSAHQCLPPFPADVMVPQYGYLTNNGSQAAGRSSF  
YCLEYFSPQMLRTGNNFQSYTTFEDVPFHSSYAHQSQSLDRMLNPLIDQLYLYLNRQTQ

TTSGTTNQSRIILFSQAGPQSMLSQARNWLPGPCYRQRLSKTANDNNNSNFPWTAASK  
YHLNGRDSLVPNPGPAMASHKDDDEKFFPMHGNLI FGKEGTTASNAELDNVMITDEEEI  
RTTNVATEQYGTVALNNQSSNTAPTTGTVNHQGLPGMWQDRDVLQGPWIWAKIPH  
TDGHFHPSLMGFGFLKHPPPPQIMIKNTVPVNPPTTFSAPKPAFSLTOYSTQGVSV  
IEWELKENSKRWNPEIOYTSYNKLSVNVDFTVDTNGVYSEPRPLCTRYLTRNL"

| BASE COUNT | 1200 a | 1295 c | 1251 g | 980 t |
|------------|--------|--------|--------|-------|
| ORIGIN     |        |        |        |       |

**Alignment Scores:**

|                        |           |               |      |
|------------------------|-----------|---------------|------|
| Pred. No.:             | 2.19e-209 | Length:       | 4726 |
| Score:                 | 3494.00   | Matches:      | 637  |
| Percent Similarity:    | 91.72%    | Conservative: | 39   |
| Best Local Similarity: | 86.43%    | Mismatches:   | 59   |
| Query Match:           | 87.59%    | Indels:       | 2    |
| DB:                    | 14        | Gaps:         | 2    |

US-09-807-802A-13 (1-736) x AVU48704 (1-4726)

|    |      |  |      |
|----|------|--|------|
| Qy | 1    | MetAlaAlaAspGlyTyrLeuProAspTrrLeuGluAspAsnLeuSerGluGlyLeuArg     | 20   |
| Db | 2209 | ATGGCTGCTGACGGTATCTTCCAGATTGGCTCGAGGACAACTTCTCTGAAGGCATTCTG      | 2268 |
| Qy | 21   | GluTrrTrrAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp     | 40   |
| Db | 2269 | GAGTGGTGGGCTCTGAACCTGGAGTCCCTCAACCCAAAGCGAACCAACACACAGGAC        | 2328 |
| Qy | 41   | AspGlyArgGlyLeuValLeuProGlyTrrLysTrrLeuGlyPropheAsnGlyLeuAsp     | 60   |
| Db | 2329 | AACCGTCGGGGTCTTGCTTCGGGTACAAATACCTCGGACCCGGTAAACGGACTCGAC        | 2388 |
| Qy | 61   | LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp        | 80   |
| Db | 2389 | AAAGGAGAGCCGGTCAACGAGCGGACGCGGCAGCCCTCGAACAACGACAAAGCTTACGAC     | 2448 |
| Qy | 81   | GlnGlnLeuLysAlaGlyAspAsnProTrrLeuArgTrrAsnHisAlaAspAlaGluPhe     | 100  |
| Db | 2449 | CAGCAGCTCAAGCGCGGTGACAAACCCGTACCTCAAGTACAAACACACGACGCGCGAGTTT    | 2508 |
| Qy | 101  | GlnGluArgLeuGlnGluAspTrrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln     | 120  |
| Db | 2509 | CAGGAGCGTCTTCAAGAAGATACGTCTTTTGGGGCAACCTTGGCAGAGCAGTCTTCCAG      | 2568 |
| Qy | 121  | AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysTrrAlaPro        | 140  |
| Db | 2569 | GCCAAAAGAGGATCCTTGAGCCTCTTGGTCTGGTTGAGGAAGCAGCTAAACGGCTCCT       | 2628 |
| Qy | 141  | GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSerGlyGlyGly     | 160  |
| Db | 2629 | GGAAAGAAGGGGGCTGTAGATCAGTCTCCTCAGGAACCGGACTCATCTCTGGTGTGGC       | 2688 |
| Qy | 161  | LysTrrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnTrrGlyAspSerGlu     | 180  |
| Db | 2689 | AAATCGGGCAACAGCCTGCCAGAAAAGACTAAATTTCTGGTCTGAGACTGGAGACTCAGAG    | 2748 |
| Qy | 181  | SerValProAspProGlnProLeuGlyGluProProAlaTrrProAlaAlaValGlyPro     | 200  |
| Db | 2749 | TCAGTCCCAGACCCCTCAACCTCTCGGAGAACCCACGAGCAGCCCCCAAGTTTGGGATCT     | 2808 |
| Qy | 201  | TrrTrrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly     | 220  |
| Db | 2809 | AATACAATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAGGGTGGCCGATGGA    | 2868 |
| Qy | 221  | ValGlyAsnAlaSerGlyAsnTrrPheHisCysAspSerTrrTrrLeuGlyAspArgValIle  | 240  |
| Db | 2869 | GTGGGTAAATTCCTCAGGAATTTGGCATTTGCCGATTTCCCAATGGCTGGGCGCAGAGTCAATC | 2928 |
| Qy | 241  | TrrTrrSerTrrArgTrrTrrAlaLeuProTrrTrrAsnAsnHisLeuTrrLysGlnIle     | 260  |
| Db | 2929 | ACCACCAGCACCCAGAACCTGGGCCCTTGCCCACTTTACAAACAACCATCTCTACAAAGCAATC | 2988 |
| Qy | 261  | SerSerAlaSerTrrGlyAlaSerAsnAspAsnHisTrrPheGlyTrrSerTrrProTrrP    | 280  |
| Db | 2989 | TCCAGCCCAATCA---GGAGCTTCAAAACGACCAACCTACTTTGGCTACAGCACCCCTGG     | 3045 |





QY 300 LeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIle 319  
Db 901 CTCATCAACAACAAATGGGGATTCCGGCCCCAAAGACTCAACTTCAAGCTGTTCAACATC 960  
QY 320 GlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSer 339  
Db 961 CAGGTCAAGGAAGTCACGACGAACGAAGGCCAACAGACCATCGCCAAATAATCTCACCAGC 1020  
QY 340 ThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHis 359  
Db 1021 ACCGTGCAGGTCCTTACGGACTCGGAGTACCAGTTACCGTACGTGCTAGGATCCGCTCAC 1080  
QY 360 GlnGlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeu 379  
Db 1081 CAGGGATGTCTGCCTCCGTTCCCGCGGACGCTCTCATGGTTCTCTCAGTACGGCTATTTA 1140  
QY 380 ThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhe 399  
Db 1141 ACTTTAAACAATGGAAGCCCAAGCCCTGGGACGTTCTCTCTACTGTCTGGAGTATTTC 1200  
QY 400 ProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluVal 419  
Db 1201 CCATCGCAGATGCTGAGAACCAGCAACACTTTCAGTTTCAGTACACCTTCGAGGACGTG 1260  
QY 420 PropheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIle 439  
Db 1261 CCTTTCCACAGCAGCTACGCGCACAGCCAGAGCTGGACAGGCTGATGAATCCCTTCATC 1320  
QY 440 AspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLys 459  
Db 1321 GACCAGTACCTGTACTACCTGGTGTGAGAACGCAACG---ACTGGAACCTGGAGGACGCAG 1377  
QY 460 AspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeu 479  
Db 1378 ACTCTGGCATTACGCCAAGCGGGTCTCTAGTCTCAATGGCCCAACCAGGCTAGAAATGGGTG 1437  
QY 480 ProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnSer 499  
Db 1438 CCCGGACCTTGCTACCGGCAGCAGCGCTCTCCACGACAACCAACCAAGAACCAACACAGC 1497  
QY 500 AsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsn 519  
Db 1498 AACTTTGGCTGGACGGAGCTGCCAAGTTTAAGTTGAACGGCCGAGACTCTCTAATGAAT 1557  
QY 520 ProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGly 539  
Db 1558 CCGGGCGTGGCAATGGCTTCCCAAGGATGACGACGACCGCTTCTTCCTTCGAGCGGG 1617  
QY 540 ValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMet 559  
Db 1618 GTCTCTGATTTTGGCAAGCAAGGAGCCGGGAACGATGGAGTGGATTACAGCCAAAGTGTG 1677  
QY 560 IleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThr 579  
Db 1678 ATTACAGATGAGGAAGAAATCAAGGCTACCAACCCGTGGCCACAGAGAATATGGAGCA 1737  
QY 580 ValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMet 599  
Db 1738 GTGGCCATCAACAACCAAGCCGCCAATACGCGAGCGCAGACCGGACTCGTGCAACAACAG 1797  
QY 600 GlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrp 619  
Db 1798 GGGGTGATTCCTCCGGCATGGTGTGGCAGAATAGACGCTGTACCTGCAGGCTCCCATCTGG 1857  
QY 620 AlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGly 639  
Db 1858 GCCAAAATTCCTCACACGACGGCAACTTTTCAACCGTCTCCCTGATGGCGGCTTTTGA 1917  
QY 640 LeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProPro 659  
Db 1918 CTGAAGCACCCCGCTCTCAAATTTCTCATCAAGAACACACCGGTTCCAGCGGACCCCGCG 1977  
QY 660 AlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnVal 679

Db 1978 CTTACCTTCAACCAAGCTGAACCTCTTTTCATCACGCAGTACAGCACCGGACAGGTC 2037  
QY 680 SerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluVal 699  
Db 2038 AGCGTGGAAATCGAGTGGGAGCTGCAGAAAAGAAAACAGCAAAACGCTGGAATCCAGAGATT 2097  
QY 700 GlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGly 719  
Db 2098 CAATACACTTCCAACACTACTACAAATCTACAAATGTGGACTTTGTCTGCTCAACACGGAGGG 2157  
QY 720 LeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736  
Db 2158 GTTTATACGAGCCTCGCCCCCATTGGCACCCGTTACCTCACCCGCAACCTG 2208  
RESULT 6  
AY243015 2217 bp DNA linear VRL 14-MAY-2003  
LOCUS Non-human primate Adeno-associated virus isolate AAVrh.10 capsid  
DEFINITION protein (VP1) gene, complete cds.  
ACCESSION AY243015  
VERSION AY243015.1 GI:29650525  
KEYWORDS  
SOURCE Non-human primate Adeno-associated virus  
ORGANISM Non-human primate Adeno-associated virus  
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.  
REFERENCE 1 (bases 1 to 2217)  
AUTHORS Gao,G., Alvira,M.R., Sonanathan,S., Lu,Y., Vandenbergh,L.H.,  
Rux,J.J., Calcedo,R., Sanniguel,J., Abbas,Z. and Wilson,J.M.  
TITLE Adeno-associated viruses undergo substantial evolution in primates  
during natural infections  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (10), 6081-6086 (2003)  
PUBMED 12716974  
REFERENCE 2 (bases 1 to 2217)  
AUTHORS Gao,G., Alvira,M.R., Sonanathan,S., Lu,Y., Vandenbergh,L.H.,  
Rux,J.J., Calcedo,R., Sanniguel,J., Abbas,Z. and Wilson,J.M.  
TITLE Direct Submission  
JOURNAL Submitted (23-FEB-2003) Department of Medicine, University of  
Pennsylvania School of Medicine and Wistar Institute, 204 Wistar  
Institute, 3601 Spruce Street, Philadelphia, PA 19104, USA  
FEATURES  
source  
1. .2217  
/organism="Non-human primate Adeno-associated virus"  
/proviral  
/mol\_type="genomic DNA"  
/isolate="AAVrh.10"  
/db\_xref="taxon:226582"  
1. .2217  
/gene="VP1"  
1. .2217  
/gene="VP1"  
/codon\_start=1  
/product="capsid protein"  
/protein\_id="AAC8201.1"  
/db\_xref="GI:29650526"  
/translation="MAADGYLPDWLEDNLSEGIREWWDLKPGAPKPKANQKQDDGRG  
LVLPGYKYLGPFGNLDKGEVNAADAALAEHDKAYDQOLKAGDNPYLRYNHADAEFQE  
RLQEDTSFGNLRGAVFQAKRVLEPLGLVEEGAKTAPGKRPVEPSPQSPDSSTGI  
GKKGQPAKKRLNFGQTGDSVPDPQPIGEPGAGPSGLSGTMAACGGAPWADNNNEG  
ADGVSSSGNWNHCDSTWLGDRVITTTSTRTWALPTYNHLYKQISNGTSGGSTNDNTYP  
GYSTPWGYFDNRFHCHFSPRDWORLNNNWGRFRKXLFNFIQVKEVTQNEGTKT  
IANNLTSTIQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMIPOGYLTLLNNSQAVGR  
SSPYCLEYFPFSQMLRTGNNFESYQFEDVPFHSSYAHYSQSLDRMLNPLIDQYLYLSR  
TQSTGGTAGTQQLLFSQAGPNNSAQAKNWLPGPCVRQQRVSTTLSQNNNSNFAWTGA  
TKYHLNGRDSLVPNGVAMATHKDDERFFPSSVLMFGKQAGKNDVDYSSVMLTSEE  
EIKTNPVATEQYGVVADNLQQNAAPIVGAVNSQALPGMWQNRDVLQGPFIWAKI  
PHTDGNFHPSPMLGFGFLKHPPPQILIKNTVPADPPTTFSSQAKLASFITQYSTGQVS  
VEIEWELQKENSKRWNPEIQYTSNYKSTNVDFAVNTDGTYSERPIGTRYLTRNL"  
BASE COUNT 563 a 673 c 565 g 416 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 6.27e-207 length: 2217

|   |         |  |     |
|---|---------|--|-----|
| Score:  | 3448.00 | Matches:   | 627 |
| Percent Similarity:                           | 91.19%  | Conservative:  | 46  |
| Best Local Similarity:                        | 84.96%  | Mismatches:  | 63  |
| Query Match:                                  | 86.44%  | Indels:  | 2   |
| DB:   | 14      | Gaps:  | 2   |
| US-09-807-802A-13 (1-736) x AY243015 (1-2217) |         |  |     |
| QY  | 1       | MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg   | 20  |
| DB  | 1       | ATGGCTGCCGATGGTTATCTTCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTCGC    | 60  |
| QY  | 21      | GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp   | 40  |
| DB  | 61      | GAGTGTGGGACTTGAAACCTGGAGCCCCGAAACCCAAAGCCAAACAGGAGAC           | 120 |
| QY  | 41      | AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp   | 60  |
| DB  | 121     | GACGGCCGGGTCTGGTGTCTTCCTGGCTACAAGTACCTCGGACCTTCAACCGGACTCGAC   | 180 |
| QY  | 61      | LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp      | 80  |
| DB  | 181     | AAGGGGAGCCCGTCAACGCGCGGACGCGGCCCTCGAGCACGACAAGGCCCTACGAC       | 240 |
| QY  | 81      | GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe   | 100 |
| DB  | 241     | CAGCAGCTCAAAGCGGGTGACAATCCGTACCTGCGGTATAACACGCGCGAGTTT         | 300 |
| QY  | 101     | GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln   | 120 |
| DB  | 301     | CAGGAGCGTCTGCAAGAAGATACGTCCTTTGGGGCAACCTCGGGGAGCAGTCTTCCAG     | 360 |
| QY  | 121     | AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro      | 140 |
| DB  | 361     | GCCAAGAAGCGGGTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGGCTAAGACGGCTCT     | 420 |
| QY  | 141     | GlyLysLysArgProValGluGlnSerProGln---GluProAspSerSerSerGlyIle   | 159 |
| DB  | 421     | GGAAGAAGAGACCGGTAGAGCCATCACCCCGAGCGTTCTCCAGACTCTCTACGGGCATC    | 480 |
| QY  | 160     | GlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSer   | 179 |
| DB  | 481     | GGCAAGAAAGGCGAGCAGCCCGCGAAAGAGACTCAACTTTGGGCGAGACTGGCGACTCA    | 540 |
| QY  | 180     | GluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGly   | 199 |
| DB  | 541     | GAGTCAGTGGCGACCCCTCAACCAATCGAGAACCCCCCGAGCCCTCTGGTCTGGGA       | 600 |
| QY  | 200     | ProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAsp      | 219 |
| DB  | 601     | TCTGGTACAATGGCTGCAGGCGGTGGCGTCCAATGGCAGACATAACGAAGGCGCCGAC     | 660 |
| QY  | 220     | GlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgVal   | 239 |
| DB  | 661     | GGAGTGGGTAGTTCCTCAGGAAATTGGCAATGGCATTCGATTCACATGGCTGGCGACAGATC | 720 |
| QY  | 240     | IleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGln   | 259 |
| DB  | 721     | ATCACCAACGACCCGAAACCTGGGCCCTCCCACTTACCAACCAACCACTCTACAAAGCAA   | 780 |
| QY  | 260     | IleSerSerAlaSerThrGlyAlaSer---AsnAspAsnHisTyrPheGlyTyrSerThr   | 278 |
| DB  | 781     | ATCTCCAACGGGACTTCGGGAGGAAGCACCAACGACAAACACCTACTTCTGGCTACAGCAC  | 840 |
| QY  | 279     | ProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGln   | 298 |
| DB  | 841     | CCCTGGGGGTATTTGACTTTAACAGATTCCACTGCCACTTCTCACACGCTGACTGGCAG    | 900 |
| QY  | 299     | ArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsn   | 318 |
| DB  | 901     | CGACTCATCAACAACAACCTGGGGATTCCGGGCCCAAGAGACTCAACTTCAAGCTCTTCAAC | 960 |
| QY  | 319     | IleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnLeuThr      | 338 |

|    |      |  |      |
|----|------|--|------|
| DB | 961  | ATCCAGGTCAGGAGGTCACGCAGAAATGAGGCACCAAGACCATCGCCATAACCTTACC       | 1020 |
| QY | 339  | SerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAla     | 358  |
| DB | 1021 | AGCACGATTACAGTCTTTACGGACTCGGAATACCAAGCTCCCGTACGTCTCGGCTCTGCG     | 1080 |
| QY | 359  | HisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyr     | 378  |
| DB | 1081 | CACCAGGGCTGCCTCCCTCCGTTCCCGGCGGACGTCTTCATGATTCTCAGTACGGGTAC      | 1140 |
| QY | 379  | LeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyr     | 398  |
| DB | 1141 | CTGACTCTGAACAATGGCAGTCAGGCCGCTGGGCCGTTCTCTCTTCTACTCTGGAGTAC      | 1200 |
| QY | 399  | PheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGlu     | 418  |
| DB | 1201 | TTTCCTTCTCAAAATGCTGAGAACGGGCAACAACCTTTGAGTTTCTAGCTACCAAGTTGAGGAC | 1260 |
| QY | 419  | ValPropheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeu     | 438  |
| DB | 1261 | GTGCCTTTTTCACAGCAGCTACGCGCACAGCCCAAGCCTGGACCGGCTGATGAACCCCTC     | 1320 |
| QY | 439  | IleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsn     | 458  |
| DB | 1321 | ATCGACCAGTACCTGTACTACCTGTCTCGGACTCAGTCCACGGGAGGTACCGCAGGAAC      | 1380 |
| QY | 459  | LysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrp     | 478  |
| DB | 1381 | CAGCAGTTGCTATTTTCTCAGGCCGGCCCTAATAACATGTCCGCTCAGGCCCAAAACTGG     | 1440 |
| QY | 479  | LeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsn     | 498  |
| DB | 1441 | CTACCCGGGCCCTGCTACCGGCAGCAACGCGTCTCCACGACACTGTTCGCAAAATAACAAC    | 1500 |
| QY | 499  | SerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIle     | 518  |
| DB | 1501 | AGCAACTTTGCTCGACCGGTGCCACCACCAAGTATCATCTGAATGGCAGAGACTCTCTGGTA   | 1560 |
| QY | 519  | AsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSer     | 538  |
| DB | 1561 | AATCCCGGTGTCGCTATGGCAACCCCAAGACGACGAGAGCGATTTTTCCTCGTCCAGC       | 1620 |
| QY | 539  | GlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnVal     | 558  |
| DB | 1621 | GGAGTCTTAATGTTTGGAAACACAGGGAGCTGGAAAAGACAACCTGGACTATAGCAGCGTT    | 1680 |
| QY | 559  | MetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGly        | 578  |
| DB | 1681 | ATGCTAACCACTGAGGAAGAAATTAACAACCAACCCAGTGGCCACAGAACAGTACGGC       | 1740 |
| QY | 579  | ThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAla     | 598  |
| DB | 1741 | GTGGTGGCCGATAACCTGCAACAGCAAAACGCCGCTCTTATTGTAGGGGCCGTCACAGT      | 1800 |
| QY | 599  | MetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIle     | 618  |
| DB | 1801 | CAAGGAGCCTTACCTGGCATGGTCTGGCAGAACCCGGGACGTGTACCTGCAGGGTCTTATC    | 1860 |
| QY | 619  | TrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPhe     | 638  |
| DB | 1861 | TGGGCCAAGATTCTCTCACAGGACGGAAACTTTTCATCTCTCCGCTGATGGGAGGCTTT      | 1920 |
| QY | 639  | GlyLeuLysAsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsnPro     | 658  |
| DB | 1921 | GGACTGAACACCCCGCTCTCTCAGATCTCTGATTAAAGATAACACCTGTTCCCGCGGATCCT   | 1980 |
| QY | 659  | ProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGln     | 678  |
| DB | 1981 | CCAACTACCTTTCAGTCAAGCTAAGCTGGCGTCTGTTTCATCAGCAGTACAGCACCCGGACAG  | 2040 |
| QY | 679  | ValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGlu     | 698  |

Db 2041 GTCAGCGTGAATTTGAATGGGAGCTGCAGAAAGAAAACAGCAACCGCTGGAACCCAGAG 2100

Qy 699 ValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsn 718  
:::|||||  
2101 ATTCAATACACTTCCAACTACTACAAATCTACAAATGTGGACTTTTGTCTTAACACAGAT 2160

Qy 719 GlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736  
|||:::|||||  
Db 2161 GGCACCTATTCTGAGCCTCGCCCCATCGGCACCGCTTACCTCACCCGTAATCTG 2214

RESULT 7  
AF513851 4721 bp DNA linear VRL 05-SEP-2002  
LOCUS Adeno-associated virus 7 nonstructural protein and capsid protein  
DEFINITION genes, complete cds.  
ACCESSION AF513851  
VERSION AF513851.1 GI:22652859  
KEYWORDS  
SOURCE Adeno-associated virus 7  
ORGANISM Adeno-associated virus 7  
REFERENCE 1 (bases 1 to 4721)  
AUTHORS Gao,G.P., Alvira,M.R., Wang,L., Calcedo,R., Johnston,J. and Wilson,J.M.  
TITLE Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (18), 11854-11859 (2002)  
PUBMED 12192090  
REFERENCE 2 (bases 1 to 4721)  
AUTHORS Alvira,M.R.  
TITLE Direct Submission  
JOURNAL Submitted (20-MAY-2002) Institute for Human Gene Therapy, University of Pennsylvania, M6.40 Maloney Bldg, 36th & Spruce Sts, Philadelphia, PA 19104, USA  
FEATURES  
source Location/Qualifiers  
1. .4721  
/organism="Adeno-associated virus 7"  
/proviral  
/mol\_type="genomic DNA"  
/specific\_host="rhesus monkey"  
/db\_xref="taxon:202812"  
334..2205  
/note="similar to AAV2 Rep 78 protein"  
/codon\_start=1  
/product="nonstructural protein"  
/protein\_id="AA03854.1"  
/db\_xref="GI:22652860"  
/translation="MFGFYEIVIKVPSDLDEHLPGISDSFVNWVAEKWELPPDSMD  
LNLIQAPLTVAEKLQDFLVQWRRVSKAPEALFFVQFEKGESYFHLHLVETTGVS  
MVLGRFLSQIREKLQVTIYRGVEPTLPNFAVTKTRNGAGGKVVDECIYPNYLLPK  
TPELQAWNTWMEYISACNLAEKRLVAQHLTHVSQTQEQNKENLNPNSDAPVIRS  
KTSARYMELVGLVDRGITSEKQWIQEDQASVISFNAASNSRQIKAAALDNAGKIMAL  
TKSAPDYLVPSPADIKTNRIYRILELNGYDPAYAGSVFLGWAQKKFKRNTIWLFG  
PATGTKNIAEATAHAPVFFYGCNVNTNENFPNDCVDKMVIWVEEGKWTAKVVESAKA  
ILGSKVRVDQCKSSAQIDPTPIVTSNTNCAVIDGNSTTFEHQQLQDRMFKFEL  
TRRLEHDFGKVTQEVKGFWRWASDHVTEVAHEFYVRKGGASKRPAPDDADISEPKRA  
CPSVADPSTSDAEGAPVDFADRYQNKCSRHAGMIQMLFPCKTCERMNQNFNICFTHGV  
RDCLECFPGVSESQPVVRKTYRKLCAIHLLGRAPEIACSDLVNVVDLDDCVSEQ"  
2222..4435  
/note="similar to AAV2 VP1 protein"  
/codon\_start=1  
/product="capsid protein"  
/protein\_id="AA03855.1"  
/db\_xref="GI:22652861"  
/translation="MAADGYLPDWLEDNLSEGIREWMDLKPGAPKPKANQKQDNDRG  
LVLPGYKYLGPFGNGLDKGEPVNAADAALEHDKAYDQQLKAGDNPYLRYNHADAEPQE  
RLQEDTSFGNLRGAVFQAKRVLPLGLVEGAKTAPAKRPVPSQPSRSPDSSTGI  
GKKQQPARKRLNFGQTGDSVPDPQLGEPAPSPSSVSGSVTAAGGAPMADNNEG  
ADGVGNAGSNWHDSTWLGDRVITTTSTRTWALPTNNHLYKQIISSETAGSTNDVTYFG  
YSTPWGYDFNRFHCHFSPRDWQRLINNHWFRPKKRLFKLFNIQVKVETNDGVTTI  
ANNLTSTIQVSDSEYQLPYVLGSAHQGCLPPFADVFMIPQYGLTLNNGSQSVGRS  
SFYCLEYFPFSQMLRTGNNEFFSYSFEDVPFHSSYADHSQMLPDQYLYYLART  
QSNPGGTAGNRELQFYQGGPSTMAEQAKNWLPGFCFRQQRVSKTLDQNNNSNFAWTGA

TKYHLNGRNSLVNPGVAMATHKDDERFFPSSGVLI FGKTGATNKTTLNVLTNTEEE  
IRPTNPVATEEYGI VSSNLQAANTAAQTQVNVNQGALPGMVQNRDYLQGP I WAKIP  
HTDGNFHPSPMLMGGFGKLP P P P Q I L I K N T P V P A N P E V F T P A K F A S F I T Q Y S T G Q V S V  
E I E W E L Q K E N S K R W N P E I Q Y T S N F E K Q T G V D F A V D S Q G V Y S E P R P I G T R Y L T R N L "

BASE COUNT 1108 a 1405 c 1301 g 907 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3.68e-206 Length: 4721  
Score: 3442.50 Matches: 630  
Percent Similarity: 92.02% Conservative: 50  
Best local Similarity: 85.25% Mismatches: 54  
Query Match: 86.30% Indels: 5  
DB: 14 Gaps: 4

US-09-807-802A-13 (1-736) x AF513851 (1-4721)

Qy 1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20  
|||  
Db 2222 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTCGC 2281  
|||

Qy 21 GluTrpTrpAspLeuLysProGlyValAlaProLysProLysAlaAsnGlnLysGlnAsp 40  
|||  
Db 2282 GAGTGGTGGACCTGAAACCTGGAGCCCGGAAACCCAAAGCCAAACAGCAAGCAGGAC 2341  
|||

Qy 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp 60  
:::|||||  
Db 2342 AACGCCCGGGTCTGGTGCTTCTCGGTACAGTACCTCGGACCTTCAACGGACTCGAC 2401  
|||

Qy 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80  
|||  
Db 2402 AAGGGGAGCCCGTCAACCGCGGAGCAGCGGCCCTCGAGACGACAAGGCCTACGAC 2461  
|||

Qy 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100  
|||  
Db 2462 CAGCAGCTCAAAGCGGGTGACAAATCCGTACCTGCGGTATAACCAACGCCGACGCGGAGTTT 2521  
|||

Qy 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120  
|||  
Db 2522 CAGGAGCGTCTGCAAGAAGATACGTCAATTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG 2581  
|||

Qy 121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro 140  
|||  
Db 2582 GCCAAGAAGCGGTTCTCGAACCTCTCGTCTGGTTGAGGAAGCGCTAAGACGGCTCCT 2641  
|||

Qy 141 GlyLysLysArgProValGluGlnSerProGln---GluProAspSerSerSerGlyIle 159  
|||  
Db 2642 GCAAAGAAGACCGGTAGAGCCGTCACCTCAGCGTTCCCCCGACTCTCCACGGGCATC 2701  
|||

Qy 160 GlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSer 179  
|||  
Db 2702 GGCAAGAAAGCCAGCAGCCCGCCAGAAAGAGACTCAATTTCGGTCAGACTGGCGACTCA 2761  
|||

Qy 180 GluSerValProAspProGlnProLeuGlyGluProAlaThrProAlaAlaValGly 199  
|||  
Db 2762 GAGTCAGTCCCGACCTCAACCTCTCGGAGAACCTCCAGACGCGCCCTCTAGTGTGGGA 2821  
|||

Qy 200 ProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnGluGlyAlaAsp 219  
|||  
Db 2822 TCTGGTACAGTGGCTGAGGCGGTGGCGCACCCTAATGGCAGACAATAACGAAGGTGCCGAC 2881  
|||

Qy 220 GlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgVal 239  
|||  
Db 2882 GGAGTGGGTAATGCCTCAGGAAATTGGCATTCGATTCACATGGCTGGCGCAGACAGATC 2941  
|||

Qy 240 IleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGln 259  
|||  
Db 2942 ATTACCACCAGCACCCGAAACCTGGGGCCCTGCCACCTACAACAACCCACTCTACAAGCAA 3001  
|||

Qy 260 IleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrPro 279  
|||  
Db 3002 ATCTCCAGTGAAGTGCAGGTAGTAGTACCACGACCAACACCTACTTCGGCTACAGCACCCCC 3061  
|||



QY 280 TrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArg 299  
Db 3062 TGGGGGTATTTTGACTTTTAACAGATTCCACTGCCACTTCTCACCACTGAGTGGCAGCGA 3121  
QY 300 LeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIle 319  
Db 3122 CTCATCAACAACAACCTGGGGATTCCGGCCCCAAGAAGCTGCGGTTCAGACTCTTCAACATC 3181  
QY 320 GlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSer 339  
Db 3182 CAGGTCAGGAGGTACGACGAATGACGGCGTTACGACCATCGCTAATAACCTTACCAGC 3241  
QY 340 ThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHis 359  
Db 3242 ACGATTGAGGTATTCTCGGACTCGGAATACCAAGTCCGCTACGCTCTCGGCTCTGCGCAC 3301  
QY 360 GlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeu 379  
Db 3302 CAGGCTGCCTGCCTCCGTTCCGGCGGACGCTTTCATGATTCTCAGTACGGCTACCTG 3361  
QY 380 ThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhe 399  
Db 3362 ACTCTCAACAATGGCAGTCAGTCTGTGGGACGTTCTCTCTTCTACTGCTGGAGTACTTC 3421  
QY 400 ProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluVal 419  
Db 3422 CCCTCTCAGATGCTGAGAACGGGCAACAACCTTGAGTTCACTACGCTTCGAGGACGTG 3481  
QY 420 ProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIle 439  
Db 3482 CCTTTCACAGCAGCTACGCACACAGCCAGAGCCTGGACCGCTGATGAATCCCTCATC 3541  
QY 440 AspGlnTyrLeuTyrTyrLeuAsnArgThrGln---AsnGlnSerGlySerAlaGlnAsn 458  
Db 3542 GACCAGTACTGTACTACCTGGCCAGAACACACAGAGTAACCCAGGAGGCACAGCTGGCAAT 3601  
QY 459 LysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrp 478  
Db 3602 CGGGAAGTGCAGTTTACCAGGGCGGGCCCTTCAACTATGGCCGAACAAGCCAAAGATTGG 3661  
QY 479 LeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsn 498  
Db 3662 TTACCTGGACCTTGCTTCGGGCAACAAGAAGTCTCCAAACGCTGGATCAAAACAACAC 3721  
QY 499 SerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIle 518  
Db 3722 AGCAACTTTGCTTGGAAGTGGTGCCCAACAATATCACCTGAACGGCAGAACTCGTTGGTT 3781  
QY 519 AsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSer 538  
Db 3782 AATCCCGGCGTCGCCATGCGCAACTCACAGGACGACGAGGACCGCTTTTCCCATCCAGC 3841  
QY 539 GlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsn---ThrAlaLeuAspAsn 557  
Db 3842 GGAGTCCTGATTTTGGAAAA-----ACTGGAGCAACTAACAAACTACATTGGAAAT 3895  
QY 558 ValMetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 577  
Db 3896 GTGTTAATGACAAATGAAGAAGAAATTGCTCTACTAATCTCTGTAGCCACGGAAGATAC 3955  
QY 578 GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis 597  
Db 3956 GGGATAGTCAGCAGCAACTTACAAGCGGCTAATACTGCAGCCCAACACAAAGTTGTCAAC 4015  
QY 598 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 617  
Db 4016 AACCAGGGAGCTTACCTGGCATGTTCTGSCAGAACCGGGACGTTGTACCTGCAGGTTCCC 4075  
QY 618 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 637  
Db 4076 ATCTGGGCCAAGATTCTCACACGGATGGCAACTTTCACCCGCTCTCTTTGATGGCGGC 4135  
QY 638 PheGlyLeuLysAsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsn 657

Db 4136 TTTGGACTTAAACATCCGCTCCTCAGATCCTGATCAAGAACACTCCCGTCCCGCTAAT 4195  
QY 658 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 677  
Db 4196 CCTCCGGAGGTGTTTACTCCTCGCAAGTTTGCTTCGTTTCATCACACAGTACAGCACCGGA 4255  
QY 678 GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro 697  
Db 4256 CAAGTCAGCGTGGAAATCGAGTGGGAGCTGCAGAAGGAAACAGCAAGCGCTGGAACCCG 4315  
QY 698 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 717  
Db 4316 GAGATTCACTACACTCCAACCTTTGAAAAGCAGACTGGTGTGGACTTTTGGCGTTGACAGC 4375  
QY 718 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736  
Db 4376 CAGGCTGTTTACTCTGAGCCTCGCCCTATTGGCACTCGTTACCTCACCCCGTAATCTG 4432  
RESULT 8  
AR264580 LOCUS AR264580 8179 bp DNA linear PAT 10-APR-2003  
DEFINITION Sequence 5 from patent US 6491907.  
ACCESSION AR264580  
VERSION AR264580.1 GI:29692850  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 8179)  
AUTHORS Rabinowitz,J.E., Samulski,R.J. and Xiao,W.  
TITLE Recombinant parvovirus vectors and method of making  
JOURNAL Patent: US 6491907-A 5 10-DEC-2002;  
FEATURES Location/Qualifiers  
source 1. 8179  
BASE COUNT 2082 a 2078 c 2027 g 1992 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.24e-205 Length: 8179  
Score: 3439.00 Matches: 626  
Percent Similarity: 90.50% Conservative: 41  
Best Local Similarity: 84.94% Mismatches: 68  
Query Match: 86.21% Indels: 2  
DB: 6 Gaps: 2  
US-09-807-802A-13 (1-736) x AR264580 (1-8179)  
QY 1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyLeArg 20  
Db 2133 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACACTCTCTCTGAAGGAATAAGA 2192  
QY 21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAlaAsnGlnLysGlnAsp 40  
Db 2193 CAGTGGTGAAGCTCAAACCTGGCCACCACCACCAAGCCCGCAGAGCGGCATAAGGAC 2252  
QY 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyPropheAsnGlyLeuAsp 60  
Db 2253 GACAGCAGGGGTCTTGTGTTCTCTGGGTACAAGTACCTCGGACCTTCAACGGGACTCGAC 2312  
QY 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80  
Db 2313 AAGGAGAGCCGGTCAACGAGGCAGACGCCCGCGGCCCTCGAGCACGACAAAGCCTACGAC 2372  
QY 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100  
Db 2373 CGGAGCTCGACAGCGGAGACAACCCGTAACCTCAAGTACAACCAACCGCCGCGGAGTTT 2432  
QY 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120  
Db 2433 CAGGAGCGCCTTAAAGAAGATACGTCTTTTGGGGGCAACCTCGGACGAGCAGTCTTCCAG 2492

|    |      |   |      |
|----|------|---|------|
| Qy | 121  | AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro       | 140  |
| Dd | 2493 | CGCAAAAAGAGGGTTCTTGAACCTCTGGGCCCTGGTTGAGGAACCTGTTAAGACGGCTCCG   | 2552 |
| Qy | 141  | GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly    | 160  |
| Dd | 2553 | GGAAAAAAGAGGCCGCTAGAGCACCTCTCTGTGGAGCGACACTCTCTCTCGGGAACCGGA    | 2612 |
| Qy | 161  | LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu    | 180  |
| Dd | 2613 | AAGCGGGCCAGCAGCTGCAAGAAAAGATTGTAATTTGGTCAGACTGGAGACGCAGAC       | 2672 |
| Qy | 181  | SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro    | 200  |
| Dd | 2673 | TCAGTACCTGACCCCGCCCTCTCGGACAGCCACCGACCCCTCTCTGGTCTGGGAACCT      | 2732 |
| Qy | 201  | ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly       | 220  |
| Dd | 2733 | AATACGATGGCTACAGCGAGTGGCGCACCAATGGCAGACAAATACGAGGGCGCCGACGGA    | 2792 |
| Qy | 221  | ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle    | 240  |
| Dd | 2793 | GTGGGTAATTCCTCCGGAATTTGGCATTTCCGATTCCCATGGATGGGCGACAGAGTCAATC   | 2852 |
| Qy | 241  | ThrThrSerThrArgThrTrpAlaLeuProThrTrpAsnAsnHisLeuTyrLysGlnIle    | 260  |
| Dd | 2853 | ACCACGACACCCGAACCTGGGCCCTGCCACCTACAAACACCACTCTACAAACAAAT        | 2912 |
| Qy | 261  | SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp    | 280  |
| Dd | 2913 | TCCAGCCCAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTGG    | 2969 |
| Qy | 281  | GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu    | 300  |
| Dd | 2970 | GGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCACCACGTGACTGGCAAGACTC     | 3029 |
| Qy | 301  | IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln    | 320  |
| Dd | 3030 | ATCAACAACNACTGGGGATTCCGACCCAGAGACTCAACTTCAAGCTCTTTAACAATCAA     | 3089 |
| Qy | 321  | ValLysGluValThrThrAsnAspGlyValThrIleAlaAsnAsnLeuThrSerThr       | 340  |
| Dd | 3090 | GTCAAAGAGGTCACGCAAGATGACGGTACGACGACGATGCGCAATTAACCTTACCAGCAGC   | 3149 |
| Qy | 341  | ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln    | 360  |
| Dd | 3150 | GTTCAGGTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTGCTCGGGTCGGCGCACCAA    | 3209 |
| Qy | 361  | GlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr    | 380  |
| Dd | 3210 | GGCTGTCTCCCGCCGTTTCCAGCGGACGCTCTTCATGGTGGTCCCTCAGTATGGATACCTCAC | 3269 |
| Qy | 381  | LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro    | 400  |
| Dd | 3270 | CTGAACAACGGAAGTCAAGCGGTGGGACGCTCATCTTTTACTGCTGGAGTACTTCCCT      | 3329 |
| Qy | 401  | SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro    | 420  |
| Dd | 3330 | TCGCAGATGCTAAGGACTGGAATAAATTCCAATTCAGCTATACCTTCGAGGATGTACCT     | 3389 |
| Qy | 421  | PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp    | 440  |
| Dd | 3390 | TTTTCACAGCAGCTACGCTCAAGCCAGAGTTTGGATCGCTTGATGAATCCTCTTATTGAT    | 3449 |
| Qy | 441  | GlnTyrLeuTyrTyrLeuAsnArgThrGln--AsnGlnSerGlySerAlaGlnAsnLys     | 459  |
| Dd | 3450 | CAGTATCTGTACTACCTGAACAGAACGCAAGGAACAACCTCTGGAACAACCAATCA        | 3509 |
| Qy | 460  | AspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeu    | 479  |
| Dd | 3510 | CGGCTGCTTTTAGCCAGGCTGGGCCCTCAGTCTATGTCTTTGCAGGCGCAGAAATTTGGCTA  | 3569 |
| Qy | 480  | ProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnSer       | 499  |

|    |      |  |      |
|----|------|--|------|
| Db | 3570 | CCTGGGCCCTGCTACCGGCAACAGAGACTTTCAAAGACTGCTAACGACAACAACACAGT    | 3629 |
| Qy | 500  | AsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsn   | 519  |
| Db | 3630 | AACCTTTCCTGGACAGCGGCCAGCAAAATATCATCTCAATGGCCGCGACTCGCTGGTGAAT  | 3689 |
| Qy | 520  | ProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGly   | 539  |
| Db | 3690 | CCAGGACCAGCTATGGCCAGTCACAAGGACGATGAAGAAAAATTTTTTCCCTATGACACGGC | 3749 |
| Qy | 540  | ValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMet   | 559  |
| Db | 3750 | AATCTAATATTGGCAAAAGAGGGACAAACGGCAAGTAACGCAGAAATTAGATAATGTAAATG | 3809 |
| Qy | 560  | IleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThr   | 579  |
| Db | 3810 | ATTACGGATGAAGAAGAGATTTCGTACCAACCAATCCTGTGGCAACAGACAGTATGGAACT  | 3869 |
| Qy | 580  | ValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMet   | 599  |
| Db | 3870 | GTGGCAAAATAACTTGCAGAGCTCAAATACAGTCCCAAGACTGGAACTGTCAATCATCAG   | 3929 |
| Qy | 600  | GlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrp   | 619  |
| Db | 3930 | GGGGCCTTACCCTGGCATGGTGTGGCAAGATCGTGACGTGTACCTTCAAGGACCTATCTGG  | 3989 |
| Qy | 620  | AlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGly   | 639  |
| Db | 3990 | GCAAAGATTCCCTCACACGGATGGACACTTTCATCCTTCTCCTCTGATGGGAGGCTTTGGA  | 4049 |
| Qy | 640  | LeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProPro      | 659  |
| Db | 4050 | CTGAACATCCGCTCCTCAAATCATGATCAAAAATACTCCGGTACCTCGGAATCCTTCG     | 4109 |
| Qy | 660  | AlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnVal   | 679  |
| Db | 4110 | ACCACCTTCAGTGGGCAAAAGTTTGCTTCTTCATCATCACAGTACTCCACGGGACAGGTC   | 4169 |
| Qy | 680  | SerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluVal   | 699  |
| Db | 4170 | AGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAAAAACAGCAACCGCTGGAATCCCGGAATT  | 4229 |
| Qy | 700  | GlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGly   | 719  |
| Db | 4230 | CAGTACACTTCCAACTACAACAAGTCTGTTAATCGTGGACTTACCGTGGATACTAATGGC   | 4289 |
| Qy | 720  | LeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu            | 736  |
| Db | 4290 | GTGTATTTCAGAGCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTG            | 4340 |

## RESULT 9

|            |   |         |     |        |                 |
|------------|---|---------|-----|--------|-----------------|
| AY243007   | AY243007  | 2217 bp | DNA | linear | VRL 14-MAY-2003 |
| LOCUS      |   |         |     |        |                 |
| DEFINITION | Non-human primate Adeno-associated virus isolate AAVrh.2 capsid protein (VP1) gene, complete cds. |         |     |        |                 |

|           |             |
|-----------|-------------|
| ACCESSION | AY243007    |
| VERSION   | AY243007.1  |
|           | GI:29650509 |

| KEYWORDS | SOURCE  | ORGANISM | REFERENCE |
|----------|---|----------|-----------|
|          | Non-human primate Adeno-associated virus                          |          |           |
|          | Non-human primate Adeno-associated virus                          |          |           |
|          | Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus. |          |           |
|          | 1 (bases 1 to 2217)   |          |           |

|         |   |
|---------|---|
| AUTHORS | Gao, G., Alvira, M.R., Somanathan, S., Lu, Y., Vandenbergh, L.H., Rux, J.J., Calcedo, R., Sanmiquel, J., Abbas, Z. and Wilson, J.M. |
| TITLE   | Adeno-associated viruses undergo substantial evolution in primates during natural infections  |

JOURNAL  
PUBMED  
Proc. Natl. Acad. Sci. U.S.A. 100 (10), 6081-6086 (2003)  
12716974

|           |   |
|-----------|---|
| REFERENCE | 2 (bases 1 to 2217)   |
| AUTHORS   | Gao, G., Alvira, M.R., Somanathan, S., Lu, Y., Vandenbergh, L.H., Rux, J., Calcedo, R., Sanmiguel, J., Abbas, Z. and Wilson, J.M. |
| TITLE     | Direct Submission   |

Submitted (23-FEB-2003) Department of Medicine, University of Pennsylvania School of Medicine and Wistar Institute, 204 Wistar Institute, 3601 Spruce Street, Philadelphia, PA 19104, USA

**FEATURES**

|        |  |
|--------|--|
| source | Location/Qualifiers  |
|        | 1..2217  |
|        | /organism="Non-human primate Adeno-associated virus"   |
|        | /proviral  |
|        | /mol_type="genomic DNA"  |
|        | /isolate="AAVrh.2"   |
|        | /db_xref="taxon:226582"  |
| gene   | 1..2217  |
|        | /gene="VP1"  |
| CDS    | 1..2217  |
|        | /gene="VP1"  |
|        | /codon_start=1   |
|        | /product="capsid protein"  |
|        | /protein_id="AAO88193.1"   |
|        | /db_xref="GI:29650510"   |
|        | /translation="MAADGYLPDWLEDNLSEGIREWWDLPKGPAPKPKANQQKQDDGRG<br>LVLPGYKYLPPFNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPVLYRNHADAEEFQ<br>RLQEDTSFGNGLGRAVFAQKRVRLEPLGLVEEGAATAFGKRPVESPQRSPDSSTGI<br>KKKHQPAKRLNFGQTGDSVPDPOPIGEPPAGPSGLSGTWAAGGAPMADNNEG<br>ADGVSSGNHWCDSITWLGDRIITSTRTWALPTNNHLYKQISNGTSGGSTNDNTYF<br>GYSTPWGFDENRFHCFSPRDQWRLLNNWGFPRKRLNFKLFNIQVKEVTQNEGTKT<br>IANNLTSTIQVFTDSEYQLPYVPSAHQGLPPFPADVFMIPOGYLTLLNNGSQAVGR<br>SSFYCLEYFQSQMLRTGNFFESYTFEDVPHSSYAHQSQSLDRLLMNPFLIDQYLYLSR<br>TQTGGTQGTQQLLFSQAGPANMSAQAKNWLPGPCYRQSTTSLSQNNNSNFAWTGA<br>TKYHLNGRDSLVPGVAMATHKDDERFPSSGVLMEFGKQAGKDNVDYSSVMLTSEE<br>EIKTTNPVATEQYGVADNLQQTNGAPIVGTVNSQGLPGMWQNRDVLQGPVWAKI<br>PHTDGNFHPSPLMGGFGLKHPPPPQILVKNTPVPADPPTTFSSQAKLASFITQYSTGQVS<br>VEIEWELQKNSKRWNPETQYTSNYYSKTNDVFAVNTGTYSEPRPIGTRYLTRNL" |

BASE COUNT 560 a 667 c 568 g 422 t

ORIGIN

Alignment Scores:

|                        |           |               |      |
|------------------------|-----------|---------------|------|
| Pred. No.:             | 2,65e-206 | Length:       | 2217 |
| Score:                 | 3438.00   | Matches:      | 624  |
| Percent Similarity:    | 91.06%    | Conservative: | 48   |
| Best Local Similarity: | 84.55%    | Mismatches:   | 64   |
| Query Match:           | 86.19%    | Indels:       | 2    |
| DB:                    | 14        | Gaps:         | 2    |

US-09-807-802A-13 (1-736) x AY243007 (1-2217)

|    |     |  |     |
|----|-----|--|-----|
| QY | 1   | MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg   | 20  |
|    |     |  |     |
| Db | 1   | ATGGCTGCCGATGGTTATCTTCAGATTGGCTTGAGGACAACTCTCTGAGGGCATTGCG     | 60  |
|    |     |  |     |
| QY | 21  | GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp   | 40  |
|    |     |  |     |
| Db | 61  | GAGTGGTGGGACCTGAACCTGGAGCCCGGAAACCCAAAGCCAAACAGCAAGCAGGAC      | 120 |
|    |     |  |     |
| QY | 41  | AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyPropheAsnGlyLeuAsp   | 60  |
|    |     |  |     |
| Db | 121 | GACGGCCGGGGTCTGGTGCTTCTCTGGCTACAAGTACTCTCGACCCCTTCAACGGACTCGAC | 180 |
|    |     |  |     |
| QY | 61  | LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp      | 80  |
|    |     |  |     |
| Db | 181 | AAGGGGGAGCCCGTCAACGCGCGGACGACGCGGCCCTCGAGCACGACGACGCGCTACGAC   | 240 |
|    |     |  |     |
| QY | 81  | GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe   | 100 |
|    |     |  |     |
| Db | 241 | CAGCAGCTCAAGCGGGTGACAAATCCGTACTCTGCGGTATAACCAACGCGCGAGTTT      | 300 |
|    |     |  |     |
| QY | 101 | GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln   | 120 |
|    |     |  |     |
| Db | 301 | CAGGACGCTCTGCAAGAAGATACGTCTTTTGGGGGCAACCTCGGGCGAGCAGCTCTCCAG   | 360 |
|    |     |  |     |
| QY | 121 | AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro      | 140 |
|    |     |  |     |
| Db | 361 | GCCAAAGAAGCGGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCGCTAAGACGGCTCCT | 420 |
|    |     |  |     |
| QY | 141 | GlyLysLysArgProValGluGlnSerProGln---GluProAspSerSerSerGlyIle   | 159 |
|    |     |  |     |



Db 1501 AGCAATTTTGCTTGGACCGGTGCCACCAAGTATCACCTGAATGGCAGAGACTCCCTGGTT 1560

Qy 519 AsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSer 538  
|||||

Db 1561 AATCCCGCGGTGGCATGGCTACCCACAAGGACGACGAGCGCTTCTTCCGCTCAAGC 1620

Qy 539 GlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnVal 558  
|||||

Db 1621 GGAGTTCTAATGTTTGGCAAGCAGGGGGCTGGAAAGACAATGTGGACTACAGCAGCGTG 1680

Qy 559 MetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGly 578  
|||||

Db 1681 ATGCTCACCAGCGAAGAGAAATTAAACTACTAACCCAGTGGCTACAGAGCAGTATGGT 1740

Qy 579 ThrValAlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAla 598  
|||||

Db 1741 GTGGTGGCAGACAACCTGCAGCAGACCAACGGAGCTCCCATTTGTGGAACTGTCAACAGC 1800

Qy 599 MetGlyAlaLeuProGlyMetValTyrGlnAspArgAspValTyrLeuGlnGlyProIle 618  
|||||

Db 1801 CAGGGGGCCTTACCTGGTATGGTCTGGCAAAACCCGGAGCTGTACCTGCAGGGGCCCATC 1860

Qy 619 TrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPhe 638  
|||||

Db 1861 TGGGCCAAAATTCCTCACACGACGGCAACTTTCATCCTTCGCCGCTGATGGAGGCTTT 1920

Qy 639 GlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnPro 658  
|||||

Db 1921 GGACTGAACACACCCGCTCCTCAGATCCTGGTGAATAACACTCCTGTTCCTGGCGATCCT 1980

Qy 659 ProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGln 678  
|||||

Db 1981 CCGACCACCTTCAGCCAGGCCAAGCTGGCTTCTTTTATCAGCAGTACAGCACCGACAG 2040

Qy 679 ValSerValGluIleGluTyrGluLeuGlnLysGluAsnSerLysArgTyrAsnProGlu 698  
|||||

Db 2041 GTCAGCTGGAATCGAATGGAGCTGCAGAAAGAAACAGAACGCTGGAACCCAGAG 2100

Qy 699 ValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 718  
:::|

Db 2101 ATTCAATATCTTCCAACTACTACAAATCTACAAATGTGGACTTTGCTGTCAATACTGAG 2160

Qy 719 GlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736  
|||

Db 2161 GGTACTATTTCAGAGCCTCGCCCATGGCACTCGTATTCTCACCCGTAATCTG 2214

RESULT 10  
AY243006 2187 bp DNA linear VRL 14-MAY-2003  
LOCUS Non-human primate Adeno-associated virus isolate AAVrh.22 capsid  
DEFINITION protein (VP1) gene, complete cds.

ACCESSION AY243006  
VERSION AY243006.1 GI:29650507  
KEYWORDS

SOURCE Non-human primate Adeno-associated virus  
ORGANISM Non-human primate Adeno-associated virus  
VIRUSES; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.

REFERENCE 1 (bases 1 to 2187)  
AUTHORS Gao,G., Alvira,M.R., Somanathan,S., Lu,Y., Vandenberghe,L.H.,  
Rux,J.J., Calcedo,R., Sanmiquel,J., Abbas,Z. and Wilson,J.M.

TITLE Adeno-associated viruses undergo substantial evolution in primates during natural infections

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (10), 6081-6086 (2003)  
PUBMED 12716974

REFERENCE 2 (bases 1 to 2187)  
AUTHORS Gao,G., Alvira,M.R., Somanathan,S., Lu,Y., Vandenberghe,L.H.,  
Rux,J., Calcedo,R., Sanmiquel,J., Abbas,Z. and Wilson,J.M.

TITLE Direct Submission  
JOURNAL Submitted (23-FEB-2003) Department of Medicine, University of Pennsylvania School of Medicine and Wistar Institute, 204 Wistar Institute, 3601 Spruce Street, Philadelphia, PA 19104, USA

FEATURES  
source 1..2187

/organism="Non-human primate Adeno-associated virus"  
/proviral  
/mol\_type="genomic DNA"  
/isolate="AAVrh.22"  
/db\_xref="taxon:226582"  
1..2187  
/gene="VP1"  
1..2187  
/gene="VP1"  
/codon\_start=1  
/product="capsid protein"  
/protein\_id="AAO88192.1"  
/db\_xref="GI:29650508"  
/translation="MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRG  
LVLPGYKYLGPFGNGLDKGEPVNAADAAALEHDKAYDQOLKAGDNPYLRYNHADAEFQE  
RLQEDTSFGGNLGRAVFOAKKRVLEPLGLVEEGAKTAPGKKRPIESPDSDSIGKKGGQ  
QPAKKLNFQGTGDSVPDPOPIGEPAPGPSGLSGTMAAGGAPMADNNEGADGVG  
NASGNWCHDSTWLGDRVITSTRTWALPTYNNHLYKQISSQSGATNDNHFFGYSTPWG  
YFDENRFCHFSPRDWQRLINNMWGFPRKLRFKLENIQVKEVTNDGVTTIANNLTS  
TIQVFSDEYQLPYVLGSAHQCLPPFPADVFMIPOYGYLTLNNGSQSVGRSSFYCLE  
YFQSMRLRTGNFEFSYTFEEVPHSSYAHQSOLDRLMNPLIDQYLYLARTQSTTGS  
TRELFHQAGPNTMAEQSKNWLPGPCYRRQRLSKDIDSNNSNFAWTGATKYHLNGRN  
SLTNPGVAMATNKDDEDOFFPINGVLVFGKTGAANKTTILENVLMTSEEEKTTNPVAT  
EEYGVSSNLQSSSTAGPOTQVNSQGALPGMVQNRDVLQGIWAKIPHTDGNFHP  
PLMGFGKLPHPPOILIKNTPVPANPPEVFTPAKFASFITQYSTGTQVSVIEIWELOKE  
NSKRWNPETQYTSYAKSNNVFAVNEGVYTEPRPIGTRYLTRNL"

BASE COUNT 556 a 667 c 560 g 404 t  
ORIGIN

Alignment Scores:  
Pred. No.: 4.64e-206 Length: 2187  
Score: 3434.00 Matches: 628  
Percent Similarity: 91.85% Conservative: 48  
Best Local Similarity: 85.33% Mismatches: 52  
Query Match: 86.09% Indels: 8  
DB: 14 Gaps: 4

US-09-807-802A-13 (1-736) x AY243006 (1-2187)

Qy 1 MetAlaAlaAspGlyTyrLeuProAspTyrLeuGluAspAsnLeuSerGluGlyIleArg 20  
|||||

Db 1 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGCGATTCCG 60  
|||||

Qy 21 GluTyrTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnLysGlnAsp 40  
|||||

Db 61 GAGTGGTGGGACTTGAAACCTGGAGCCCCGAAACCCAAAGCCACCAAGCAGGACGAG 120  
|||||

Qy 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp 60  
|||||

Db 121 GACGGCCGGGGTCTGGTGCTTCTGGCTACAGTACCTCGGACCTTCAACGCGACTCGAC 180  
|||||

Qy 61 LysGlyGluProValAsnAlaAlaAspAlaAlaAlaLeuGluHisAspLysAlaTyrAsp 80  
|||||

Db 181 AAGGAGAGCCGGTCAACGGCGGACGCGAGCGGCCCTCGAGCACGACAGGCTACGAC 240  
|||||

Qy 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100  
|||||

Db 241 CAGCAGCTCAAAGCGGTGACAAATCCGTACCTCGCGGTATAACCAAGCCGCGAGTTT 300  
|||||

Qy 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120  
|||||

Db 301 CAGGAGCGTCTTCAAGAAGATACGTCTTTTGGGGCAACCTCGGGCGAGAGTCTTCCAG 360  
|||||

Qy 121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro 140  
|||||

Db 361 GCCAAGAAGCGGGTTCTCGAACCTCTCGGTCTGGTGTGAGGAAGGCGTAAAGACGGCTCCT 420  
|||||

Qy 141 GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerGlyIleGly 160  
|||||

Db 421 GGAAAGAAGAGACCCCATAGAA-----TCCCCCGACTCTCTCCACGGGCATCGGC 468  
|||||

Qy 161 LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180  
|||

Db 469 AAGAAAGCCAGCAGCCCGCTAAAAAGAAGCTCAACTTTGGCGAGACTGGCGACTCAGAG 528

QY 181 SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro 200

Db 529 TCAGTGCCTGACCTCAACCAATCGGAGAACCCCGCAGGCCCTCTGGTCTGGGATCT 588

QY 201 ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly 220

Db 589 GGTACAATGGCTGCAGCGGTGGCGCTCCAATGGCAGACAAATAACGAAGCGCCGACGGA 648

QY 221 ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle 240

Db 649 GTGGTAATGCCTCCGGAATTTGGCATTTGGATTCCACATGGCTGGCGACAGATCATC 708

QY 241 ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle 260

Db 709 ACCACGACCCCGCACCTGGGCCCTGCCACCTACAAACACCACTCTACAAAGCAGATA 768

QY 261 SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp 280

Db 769 TCAAGTCAGAGC---GGGGCTACCAACGACAACCACTTCTTCGGCTACAGCACCCCTCGG 825

QY 281 GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu 300

Db 826 GSCTATTTTGACTTCAACAGATTCCACTGCCACTTCTCACCGAGTGACTGGCAGCGACTC 885

QY 301 IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln 320

Db 886 ATCAACAACAACACTGGGATTCCGGCCCGAGAAAGCTCGGGTTCAAGTTGTTCAACATCCAG 945

QY 321 VallysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThr 340

Db 946 GTCAAGGAGGTTCAGACGAAACGACGGCGTTACGACCATCGCTAATAACCTTACCAGCAG 1005

QY 341 ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln 360

Db 1006 ATTCAGGTCTTCTCGGACTCGGAGTACCAACTGCCGTACGTCCTCGGCTCTGCGCACAG 1065

QY 361 GlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr 380

Db 1066 GGCTGCCTCCCTCCGTTCCCTGCGGACGTGTTTCATGATTCCTCAGTACGGATATCTGACT 1125

QY 381 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 400

Db 1126 CTAACAACGCGCAGTCAGTCTGTGGGAGTTCCTCTCTTCTACTGCTGGAGTACTTTCTCT 1185

QY 401 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro 420

Db 1186 TCTCAGATGTGAGAACGGGCAATAACTTTGAATTTCAGCTACACCTTTGAGGAAGTGCT 1245

QY 421 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 440

Db 1246 TTCCACAGCAGCTATGGCCACAGCCAGAGCCTGGACCGGTGATGAATCCCTCATCGAC 1305

QY 441 GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAsp 460

Db 1306 CAGTACCTGTACTACCTGGCCCGGACCCAGAGCACTACGGGGTCC-----ACAAGGGAG 1359

QY 461 LeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuPro 480

Db 1360 CTGCAGTTCCATCAGGCTGGGCCCAACACCATGGCCGAGCAATCAAGAAGACTGGTGCCC 1419

QY 481 GlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsn 500

Db 1420 GGACCCCTGTTATCGCGCGCAGAGACTGTCAAAAGACATAGACAGCAACAACAACAGTAAC 1479

QY 501 PheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnPro 520

Db 1480 TTTGCCTGGACCGGGCCCACTAAATACCATCTGAATGGTAGAAATTCATTAAACCAACCCG 1539

QY 521 GlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyVal 540

Db 1540 GCGGTAGCCATGGCCACCACCAAGGACGACGAGGACCAGTCTCTTCTCCCATCAACGGAGTG 1599

QY 541 MetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIle 560

Db 1600 CTGGTTTTTGGCAAAACGGGGGCT---GCCAACAGACAACGCTGGAAAACGTTGCTAATG 1656

QY 561 ThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrVal 580

Db 1657 ACCAGCGAGGAGAGATCAAAACCAACCACTCCGTTGGCTACAGAAGATACCGTGTGGTC 1716

QY 581 AlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGly 600

Db 1717 TCCAGCAACCTGCAATCGTCTACGGCCGGACCCAGACACAGACTGTCAACAGCCAGGGG 1776

QY 601 AlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAla 620

Db 1777 GCTCTGCCCGCATGCTTGGCAGAACCCGGAGCGTGACCTGCAGGTCCTCATCTGGGCC 1836

QY 621 LysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeu 640

Db 1837 AAAATTCTCACAGGACGGCAACTTTCACCCGCTCTCCCTGTATGGCGGATTTGGACTC 1896

QY 641 LysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAla 660

Db 1897 AAACACCCGCCCTCTCAAATTCTCATCAAAACACCCCGGTACCTGCTAATCTCCAGAG 1956

QY 661 GluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSer 680

Db 1957 GTGTTTACTCCTGCCAAGTTTGCTTCATTATCAGCAGTACAGCACCCGCCAGGTGAGC 2016

QY 681 ValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGln 700

Db 2017 GTGGAGATCGAGTGGGAACGTGAGAAAGAGAAACAGCAACCGTGGAAATCCAGAGATT 2076

QY 701 TyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeu 720

Db 2077 TACACCTCAAATTATGCAAGTCTAATAATGTGGAATTTGCTGTCAACAACGAAGGGGTT 2136

QY 721 TyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736

Db 2137 TATACTGAGCCTCGCCCATTTGGCACCCCGTTACTTCACCCGTAACCTG 2184

RESULT 11

AY243009  
LOCUS AY243009 2217 bp DNA linear VRL 14-MAY-2003  
DEFINITION Non-human primate Adeno-associated virus isolate AAVrh.18 capsid protein (VP1) gene, complete cds.  
ACCESSION AY243009  
VERSION AY243009.1 GI:29650513  
SOURCE Non-human primate Adeno-associated virus  
ORGANISM Non-human primate Adeno-associated virus  
REFERENCE 1 (bases 1 to 2217)  
AUTHORS Gao,G., Alvira,M.R., Somanathan,S., Lu,Y., Vandenbergh,L.H., Rux,J.J., Calcedo,R., Sanmiguel,J., Abbas,Z. and Wilson,J.M.  
TITLE Adeno-associated viruses undergo subantial evolution in primates during natural infections  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (10), 6081-6086 (2003)  
PUBMED 12716974  
REFERENCE 2 (bases 1 to 2217)  
AUTHORS Gao,G., Alvira,M.R., Somanathan,S., Lu,Y., Vandenbergh,L.H., Rux,J., Calcedo,R., Sanmiguel,J., Abbas,Z. and Wilson,J.M.  
TITLE Direct Submision  
JOURNAL Submitted (23-FEB-2003) Department of Medicine, University of Pennsylvania School of Medicine and Wistar Institute, 204 Wistar Institute, 3601 Spruce Street, Philadelphia, PA 19104, USA  
FEATURES  
source Location/Qualifiers  
1. .2217  
/organism="Non-human primate Adeno-associated virus"  
/proviral  
/mol\_type="genomic DNA"  
/isolate="AAVrh.18"  
/db\_xref="taxon:226582"

|   |   |  |       |       |  |
|---|---|--|-------|-------|--|
| gene  | 1. .2217  |  |       |       |  |
|   | /gene="VP1"   |  |       |       |  |
| CDS   | 1. .2217  |  |       |       |  |
|   | /gene="VP1"   |  |       |       |  |
|   | /codon start=1  |  |       |       |  |
|   | /product="capsid protein"                                   |  |       |       |  |
|   | /protein_id="AAO88195.1"                                    |  |       |       |  |
|   | /db_xref="GI:29650514"                                      |  |       |       |  |
|   | /translation="MAADGYLPDWLEDNLSEGIREWWDLPKPGAKPKANQQKQDDGRG  |  |       |       |  |
|   | LVLPGYKYLGPFGNGLDKGEVNEADAALEHDKAYDKOLEQDNFYLKYNHADAEEFQE   |  |       |       |  |
|   | RLQEDTSFGGNLGRAVFOAKKRVLEPLGLVEEGAKTAPGKRVPESQRSDDSTGI      |  |       |       |  |
|   | GKTQQPAKKRLNFGQTDSESVDPDPPIGEPPAGPSGLSGTMAAGGGAPMADNNEG     |  |       |       |  |
|   | ADGVSSSGNWHCDSTLWGLDRVITTSRTWALPTYNNHLYKQISNGTSGGSTNDNTYF   |  |       |       |  |
|   | GYSTPWGYFDNRFHCHFSPRDWORLNNWGRPKRLNFKLFNIQVKEVTQNEGKTK      |  |       |       |  |
|   | IANNLTSTIQFTDSEYQLPYVLGSAHQGLPPFPADVFMIPQYGLTLNNGSQAVGR     |  |       |       |  |
|   | SSPYCLEYFPQSMLRTGNPFESYQFEDVPFHSSYAHSQSLDRIMNPLIDQYLYLSR    |  |       |       |  |
|   | TQSTGGTAGTQQLFSGAPNNMSAQAKNWLPGPCYRQQRVSTLTSQNNNSNFAWTGA    |  |       |       |  |
|   | TKYHLNGRDSLVNPGVAMATHKDDDEERFFPSSGVLFMFGKQAGKDNVDYSSVMLTSEE |  |       |       |  |
|   | EIKTNPVATEQYGVVADNLQQNAAPIVGAVNSQGLALPGMVQNRDVLQGPPIWAKI    |  |       |       |  |
|   | PHTDGNFHPSPLMGGFGLKHPPQILIKNTVPVADPPTTFESQAKLASFITQYSTGQVS  |  |       |       |  |
|   | VEIEWELQKENSKRWNPEIQYTSNYKYKSTNVDFAVNTEGTYSEPRPIGTRYLTRNL"  |  |       |       |  |
| BASE COUNT                                    | 563 a   | 673 c  | 568 g | 413 t |  |
| ORIGIN  |   |  |       |       |  |
| Alignment Scores:                             |   |  |       |       |  |
| Pred. No.:                                    | 6.3e-206  | Length:  | 2217  |       |  |
| Score:  | 3432.00   | Matches:   | 623   |       |  |
| Percent Similarity:                           | 90.92%  | Conservative:  | 48    |       |  |
| Best Local Similarity:                        | 84.42%  | Mismatches:  | 65    |       |  |
| Query Match:                                  | 86.04%  | Indels:  | 2     |       |  |
| DB:   | 14  | Gaps:  | 2     |       |  |
| US-09-807-802A-13 (1-736) x AY243009 (1-2217) |   |  |       |       |  |
| Qy  | 1   | MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg   | 20    |       |  |
| Db  | 1   | ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC    | 60    |       |  |
| Qy  | 21  | GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnLysGlnAsp      | 40    |       |  |
| Db  | 61  | GAGTGTGGGACTTGAAACCTGGAGCCCCGAAACCCAAAGCCACAGCAAGGACGAGGAC     | 120   |       |  |
| Qy  | 41  | AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp   | 60    |       |  |
| Db  | 121   | GACGCGCGGGTCTGGTGCTTCTCTGGCTACAAGTACCTCGGACCCCTTCAACGGACTCGAC  | 180   |       |  |
| Qy  | 61  | LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp      | 80    |       |  |
| Db  | 181   | AAGGGAGAGCCGGTCAACGAGGACAGCGCCGCGGCCCTCGAGCACGACAAGGCCCTACGAC  | 240   |       |  |
| Qy  | 81  | GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe   | 100   |       |  |
| Db  | 241   | AAGCAGCTCGAGCAGGGGGACAAACCCGTACCTCAAGTACAACCAACGCGCGACGCGAGTTT | 300   |       |  |
| Qy  | 101   | GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln   | 120   |       |  |
| Db  | 301   | CAGGAGCGTCTTCAAGAAGATACGTCTTTTGGGGCAACCTCGGGCGAGCAGTCTTCCAG    | 360   |       |  |
| Qy  | 121   | AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro      | 140   |       |  |
| Db  | 361   | GCCAAGAAGCGGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGCTCCT    | 420   |       |  |
| Qy  | 141   | GlyLysLysArgProValGluGlnSerProGln--GluProAspSerSerSerGlyIle    | 159   |       |  |
| Db  | 421   | GGAAAGAAGAGACCGGTAGAGCCATCACCCAGCGTTTCTCCAGACTCCTCTACGGGCATC   | 480   |       |  |
| Qy  | 160   | GlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSer   | 179   |       |  |
| Db  | 481   | GGCAAGACAGGCCAGCAGCCCCGCAAAAGAGACTCAACTTTGGGCAGACTGGCGACTCA    | 540   |       |  |
| Qy  | 180   | GluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGly   | 199   |       |  |
| Db  | 541   | GAGTCAGTCCCGACCCCTCAACCAATCGGAGAACCCCCCGAGGCCCTCTGGTCTGGGA     | 600   |       |  |

|    |      |   |      |  |
|----|------|---|------|--|
| QY | 200  | ProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAsp       | 219  |  |
| Db | 601  | TCTGGTACAATGGCTGCAGGCGGTGGCGCTCCAATGGCAGACAATAACGAAGCGCCGAC     | 660  |  |
| QY | 220  | GlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgVal    | 239  |  |
| Db | 661  | GGAGTGGGTAGTTCTCTCAGGAAATTGGCATTTGGCATTCACATGGTGGCGACAGAGTC     | 720  |  |
| QY | 240  | IleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGln    | 259  |  |
| Db | 721  | ATCACCACCAAGCACCGGAACCTGGGCCCTCCACACCTACAACAACCACTCTACAAGCAA    | 780  |  |
| QY | 260  | IleSerSerAlaSerThrGlyAlaSer---AsnAspAsnHisTyrPheGlyTyrSerThr    | 278  |  |
| Db | 781  | ATCTCCAACGGGACATCGGAGGAAGCACCAACGACAAACACCTACTTCGGCTACAGCACC    | 840  |  |
| QY | 279  | ProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGln    | 298  |  |
| Db | 841  | CCCTGGGGGTATTTTGAATTTAACAGATTCCACTGCCACTTCTCACCACGTGACTGGCAG    | 900  |  |
| QY | 299  | ArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsn    | 318  |  |
| Db | 901  | CGACTCATCAACAACAACTGGGGATTCCGGCCCAAGAGACTCAACTTCAAGCTCTTCAAC    | 960  |  |
| QY | 319  | IleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThr    | 338  |  |
| Db | 961  | ATCCAGGTCAAGGAGGTACGCAGAAATGAAGGCCAACCAAGACCATCGCCAATAACCTTACC  | 1020 |  |
| QY | 339  | SerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAla    | 358  |  |
| Db | 1021 | AGCAGATTACAGTCTTTACGGACTCGGAATACCAGCTCCCGTACGCTCCTCGCTCTGCG     | 1080 |  |
| QY | 359  | HisGlnGlyCysLeuProPheProPheProAlaAspValPheMetIleProGlnTyrGlyTyr | 378  |  |
| Db | 1081 | CACCAGGGCTGCCTCCGCTCCGTCGGGGGAGCGTCTTCATGATTCTCAGTACGGGTAC      | 1140 |  |
| QY | 379  | LeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyr    | 398  |  |
| Db | 1141 | CTGACTCTGAACAACCGCAGTCAGGCCGTGGGCGTTCCTCTTCTACTGCCCTGGAGTAC     | 1200 |  |
| QY | 399  | PheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGlu    | 418  |  |
| Db | 1201 | TTTCTCTTCTCAAAATGCTGAGAACCGGGCAACAACATTTGAGTTTCCAGTACCAAGTTT    | 1260 |  |
| QY | 419  | ValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeu    | 438  |  |
| Db | 1261 | GTGCTTTTTCACAGCAGCTACGCGCACAGCCAAAGCCTGGACCGGCTGATGAACCCCTC     | 1320 |  |
| QY | 439  | IleAspGlnTyrLeuTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsn       | 458  |  |
| Db | 1321 | ATGACCAAGTACCTGTACTACCTGTCTCGGACTCAGTCCACGGGAGGTACCGCAGGAAC     | 1380 |  |
| QY | 459  | LysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrp    | 478  |  |
| Db | 1381 | CAGCAGTTGGTATTTTCTCAGGCCGGCCCTAATAACATGTCCGGCTCAGGCCCAAAACTGG   | 1440 |  |
| QY | 479  | LeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsn    | 498  |  |
| Db | 1441 | CTACCCGGGCCCTGCTACCGGCAGCAACGCGCTCTCCACGACACTGTCCGCAAAATAACAAC  | 1500 |  |
| QY | 499  | SerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIle    | 518  |  |
| Db | 1501 | AGCAACTTTGCTTGGACCGGTGCCACCAAGTATCATCTGAATGGCAGAGACTCTCTGGTA    | 1560 |  |
| QY | 519  | AsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSer    | 538  |  |
| Db | 1561 | AATCCCGGTGTCGTATGGCAACGCAACAGGACGACGAAGAGCGATTTTTTCCATCCAGC     | 1620 |  |
| QY | 539  | GlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnVal    | 558  |  |
| Db | 1621 | GGAGTCTTGATGTTTGGGAAACAGGGAGCTGGGAAAGACAACGTTGGACTATAGCAGCGTT   | 1680 |  |





QY 220 GlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgVal 239  
Db 661 GGAGTGGGTAGTTCTCTCAGGAATTGGCATTCGATTCACATGGCTGGCGCAGAGTC 720  
QY 240 IleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGln 259  
Db 721 ATCACCACAGCACCCGAACTGGGCCCTCCCACTTACAACAACCACTTACAAGCAA 780  
QY 260 IleSerSerAlaSerThrGlyAlaSer---AsnAspAsnHisTyrPheGlyTyrSerThr 278  
Db 781 ATCTCCAACGGGACTTCGGGGAAGCACCAACGACAACACCTACTTCGGCTACAGCACC 840  
QY 279 ProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGln 298  
Db 841 CCCTGGGGGTATTTTGACTTTAAACAGATTCCACTGCCACTTCTCACCACGTCGACTGGCAG 900  
QY 299 ArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsn 318  
Db 901 CGACTCATCAACAACAACACTGGGGATTCCGGGCCCAAGAGACTCAACTTCAAGCTCTTCAAC 960  
QY 319 IleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThr 338  
Db 961 ATCCAGGTCAAGGAGGTTCAGCAGAAATGAAGGCACCAAGACCATCGCCAATAACCTTACC 1020  
QY 339 SerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAla 358  
Db 1021 AGCAGGATTCAGGTCTTTACGGACTCGGAATACCAAGTCCCGTACGTCCTCGGCTCTGCG 1080  
QY 359 HisGlnGlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGlyTyr 378  
Db 1081 CGCCAGGGCTGCCTGCCTCCGTTCCCGCGGACGCTTTCATGATTCTCAGTACGGGTAC 1140  
QY 379 LeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyr 398  
Db 1141 CTGACTCTGAACAATGGCAGTCAGGCGCGTGGGCGGTTCTCTCTTACTGCTGGAGTAC 1200  
QY 399 PheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGlu 418  
Db 1201 TTTCTCTCTCAAATGCTGAGAACGGGCAACAACCTTTTGAGTTCAAGTACCAGTTTGAGGAC 1260  
QY 419 ValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeu 438  
Db 1261 GTGCCTTTTCACAGCAGCTACGCGCACAGCCAAAGCCTGGACCGGTGATGAACCCCTC 1320  
QY 439 IleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsn 458  
Db 1321 ATCGACCAGTACCTGTACTACCTGTCTCGGACTCAGTCCAGGGAGGTACCCGAGGAAC 1380  
QY 459 LysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrp 478  
Db 1381 CAGCAGTTGCTATTTTCTCAGGCGGGCCCTAATAACATGTCGGCTCAGGCCAAAACTGG 1440  
QY 479 LeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 498  
Db 1441 CTACCCGGGCCCTGTCTACCGGCAGCAACCGCGTCTCCACGACACTGTCCGCAAAATAACAAC 1500  
QY 499 SerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIle 518  
Db 1501 AGCAACTTTGCTGGACCGGTGCGCAACCAAGTATCATCTGAATGGCAGAGACTCTCTGTA 1560  
QY 519 AsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSer 538  
Db 1561 AATCCCGGTGTCGCTATGGCAACCCACAGGACGACGAGAGCGATTTTTCCTCCAGC 1620  
QY 539 GlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnVal 558  
Db 1621 GGAGTCTTAATGTTTGGGAAACAGGGAGCTGGAAAAGGCAACGTGGACTATAGACGCTT 1680  
QY 559 MetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGly 578  
Db 1681 ATGCTAACCAAGTGGAGGAATAATAAACCACCAACCCAGTGGCCACAGAACAGTACGGC 1740  
QY 579 ThrValAlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAla 598

Db 1741 GTGGTGGCCGATAACTGTCAACAGCAAAACGCCGCTCCTATTGTAGGGCCGTCAACAGT 1800  
QY 599 MetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIle 618  
Db 1801 CAAGGAGCCTTACCTGGCATGGTCTGGCAGAAACCGGACGTGTACCTCGAGGGTCTATC 1860  
QY 619 TrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPhe 638  
Db 1861 TGGGCCAAGATTCTCACACGACGCGGAAACTTTCATCCCTCGCCGCTGATGGGGGCTTT 1920  
QY 639 GlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnPro 658  
Db 1921 GGACTGAACACCCGCCCTCTCTCAGATCTCTGATTAGAATACACCTGTTCGCGGATCCT 1980  
QY 659 ProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGln 678  
Db 1981 CCAACTACCTTCAGTCAAGCTAAGCTGGCGTCTTCATCAGCAGTACAGCACCGGACAG 2040  
QY 679 ValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGlu 698  
Db 2041 GTCAGCGTGGAATTTGAATGGGAGCTGCAGAAAGAAACAGCAACGCTGGAACCCAGAG 2100  
QY 699 ValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 718  
Db 2101 ATTCAATACACTTCCAACACTACTACAAATCTACAAATGTGGACTTTTGTGTAAACACAGAT 2160  
QY 719 GlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736  
Db 2161 GGCACTTATTCTGAGCCTCGCCCATCGGCACCCGTTACCTCACCCGTAATCTG 2214  
RESULT 13  
AY243020 2214 bp DNA linear VRL 14-MAY-2003  
LOCUS Non-human primate Adeno-associated virus isolate AAVcy.2 capsid  
DEFINITION protein (VP1) gene, complete cds.  
ACCESSION AY243020  
VERSION AY243020.1 GI:29650535  
KEYWORDS  
SOURCE Non-human primate Adeno-associated virus  
ORGANISM Non-human primate Adeno-associated virus  
REFERENCE Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.  
1 (bases 1 to 2214)  
AUTHORS Gao,G., Alvira,M.R., Somanathan,S., Lu,Y., Vandenberghe,L.H.,  
Rux,J.J., Calcedo,R., Sanmiguell,J., Abbas,Z. and Wilson,J.M..  
TITLE Adeno-associated viruses undergo substantial evolution in primates  
during natural infections  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (10), 6081-6086 (2003)  
PUBMED 12716974  
REFERENCE 2 (bases 1 to 2214)  
AUTHORS Gao,G., Alvira,M.R., Somanathan,S., Lu,Y., Vandenberghe,L.H.,  
Rux,J., Calcedo,R., Sanmiguell,J., Abbas,Z. and Wilson,J.M..  
TITLE Direct Submision  
JOURNAL Submitted (23-FEB-2003) Department of Medicine, University of  
Pennsylvania School of Medicine and Wistar Institute, 204 Wistar  
Institute, 3601 Spruce Street, Philadelphia, PA 19104, USA  
FEATURES  
source Location/Qualifiers  
1. .2214  
/organism="Non-human primate Adeno-associated virus"  
/proviral  
/mol\_type="genomic DNA"  
/isolate="AAVcy.2"  
/db\_xref="taxon:226582"  
1. .2214  
/gene="VP1"  
1. .2214  
/gene="VP1"  
/codon\_start=1  
/product="capsid protein"  
/protein\_id="AAO88206.1"  
/db\_xref="GI:29650536"  
/translation="MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDNRG  
LVLPGYKYLGPFGNLDKGPVNAADAAALEHDKAYDQQLNAGDNPYLRYNHADAEFQE

RLQEDTSFGGNLGRAVFOAKKRVLEPLGLVEEGAKTAPAKKRPVPSPOKSPDSSTGI  
GKKGQQPARKRLNFGQTDSESVDPDPQLGEPAPSSVSGTVAAGGAPMADNNEG  
ADGVGNASGNWHCDSTWLGDVITTSRTWALPTVNNHLYEQIISSETAGSTNDNTYFG  
YSTPWGYFDENRFHCFSPRDWQRLINNNWGFPRPKLRFKLFNIQVKEVTNDGVTTI  
ANNLTSTIQVFDSEYQLPVVLGSAHQGCLPPFPADVFMIPOYGYLTLLNQSQSVGRS  
SPYCLEYFPQSOMLRTGNPFYSFEDVPFHSSYAHQSLSRLMNPPLIDQYLYLART  
QSDPGGTAGNRELQFYQGGPSTMAEQAKNWLPGPCFRQORVSKTLDQNNNSNFAWTGA  
TKYHLNGRNSLVNPGVAMATHKDEDFPSPSSVLIFFKGTGATNKTTLENVLMTNEEE  
IRPTNPVATEEYGIIVSSNLQAANTAAQTQVNNQGLPGMWQNRDVLQCPHWAKIP  
HDTGNFHPSPLMGGFGLKHPPPQILIKNTVPNPNPEVFTPAKFASFITQYSTQVSV  
EIEWELQKENSKRWDPEIQYTSNFEKQTGVDFAVDSQGVSEPRPIGTRYLTRNL"

BASE COUNT 557 a 663 c 565 g 429 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.86e-205 Length: 2214  
Score: 3424.50 Matches: 626  
Percent Similarity: 91.88% Conservative: 53  
Best Local Similarity: 84.71% Mismatches: 55  
Query Match: 85.85% Indels: 5  
DB: 14 Gaps: 4

US-09-807-802A-13 (1-736) x AY243020 (1-2214)

|    |     |  |     |
|----|-----|--|-----|
| QY | 1   | MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg | 20  |
| Db | 1   | ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTCCG | 60  |
| QY | 21  | GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp | 40  |
| Db | 61  | GAGTGGTGGACCTGAACCTGGAGCCCGGAAACCCAAAGCCCAACGAGCAAGGAGAC     | 120 |
| QY | 41  | AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp | 60  |
| Db | 121 | AACGGCCGGGTCTGGTGTCTTCTGGCTACAAGTACCTCGGACCTTCAACGGACTCGAC   | 180 |
| QY | 61  | LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp    | 80  |
| Db | 181 | AAGGGGAGCGCGTCAACGCGGACGCGGCGCCCTCGAGCAGCAAGGCCTACGAC        | 240 |
| QY | 81  | GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe | 100 |
| Db | 241 | CAGCAGCTCAATGCGGGTGACAACCCGTACCTGCGGTATACCAACGCGCAGCGGAGTTT  | 300 |
| QY | 101 | GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln | 120 |
| Db | 301 | CAGGAGCGTCTGCAAGAAGATACGTCAATTGGGGGCAACCTCGGCGAGCAGTCTTCCAG  | 360 |
| QY | 121 | AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro    | 140 |
| Db | 361 | GCCAAGAAGCGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGSCGTAAACGCGTCTCT   | 420 |
| QY | 141 | GlyLysLysArgProValGluGlnSerProGln---GluProAspSerSerSerGlyIle | 159 |
| Db | 421 | GCAAGAAGAGACCGGTAGAGCCGTCACTCAGCGTTCCCGGACTCTCTCCACGGGCATC   | 480 |
| QY | 160 | GlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSer | 179 |
| Db | 481 | GGCAAGAAGGCCAGCAGCCCGCCAGAAAGAGACTCAATTTCCGTCAGACTGGCGACTCA  | 540 |
| QY | 180 | GluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGly | 199 |
| Db | 541 | GAGTCAGTCCCGACCTCAACCTCTCGGAGAACCTCCAGCAGCGCCCTCTAGTGTGGGA   | 600 |
| QY | 200 | ProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAsp    | 219 |
| Db | 601 | TCTGGTACAGTGGCTGCAGCGGTGGCGCACCAATGGCAGACAATAACGAAGGTGCCGAC  | 660 |
| QY | 220 | GlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgVal | 239 |
| Db | 661 | GGAGTGGGTAATGCCTCAGGAAATTGGCATTCGATTCCACATGGCTGGCGGACAGATC   | 720 |
| QY | 240 | IleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGln | 259 |

|    |      |  |      |
|----|------|--|------|
| Db | 721  | ATTACCACGACCCGAAACCTGGGCCCTGCCACCTACAACAACCACTCTACGAGCAA     | 780  |
| QY | 260  | IleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrPro | 279  |
| Db | 781  | ATCTCCAGTAAACTGCAGGTAGTACCAACGACGAACACCTACTTCGGCTACAGCACCC   | 840  |
| QY | 280  | TrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArg | 299  |
| Db | 841  | TGGGGGTATTTTGACTTTAAACAGATTCCACTGCCACTTCTCACACGCTGACTGGCAG   | 900  |
| QY | 300  | LeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIle | 319  |
| Db | 901  | CTCATCAACAACAACACTGGGGATTCCGGGCCCAAGAAGCTGCGGTTCAAGCTCTTCA   | 960  |
| QY | 320  | GlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSer | 339  |
| Db | 961  | CAGGTCAAGGAGGTACGACGAATGACGGCGTTACGACCATCGCTAATAACCTTACCAG   | 1020 |
| QY | 340  | ThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHis | 359  |
| Db | 1021 | ACGATTCAGGTATTTCTCGGACTCGGAATACCAAGTCCGTACGTCTCGGCTCTGGCAC   | 1080 |
| QY | 360  | GlnGlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeu    | 379  |
| Db | 1081 | CAGGGCTGCCTCCGTTCCCGGGGACGTTCTCATGATTCTCTCAGTACGGTACCTG      | 1140 |
| QY | 380  | ThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhe | 399  |
| Db | 1141 | ACTCTCAACAATGCGAGTCAGTCTGTGGACGTTCTCTTACTGCTCGGAGTACTTC      | 1200 |
| QY | 400  | ProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluVal    | 419  |
| Db | 1201 | CCCTCTCAGATGCTGAGAAACGGGCAACAACCTTTGAGTTTCTGAGTTCAGCTACAG    | 1260 |
| QY | 420  | ProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIle | 439  |
| Db | 1261 | CCTTTCACAGCAGCTACGCACACAGCCAGAGCCTGGACCGGCTGATGAATCCCTCAT    | 1320 |
| QY | 440  | AspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGln---SerGlySerAlaGlnAsn | 458  |
| Db | 1321 | GACCAGTACTTGTACTACCTGGCCAGAAACACAGAGTGACCCAGGAGGCACAGCTGG    | 1380 |
| QY | 459  | LysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrp | 478  |
| Db | 1381 | CGGGAACCTGCAGTTTACAGGGCGGGCCCTTCAACTATGGCCGAACAAGCCCAAGAT    | 1440 |
| QY | 479  | LeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsn | 498  |
| Db | 1441 | TTACCTGGACCCCTGCTTCCGGCAACAAGGGTCTCCAAAACGCTGGATCAAAAACAAC   | 1500 |
| QY | 499  | SerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIle | 518  |
| Db | 1501 | AGCAACTTTGCTTGGACTGGTGCCACCAATATACCTGAACGGCAGAAACTCGTTGGTT   | 1560 |
| QY | 519  | AsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSer | 538  |
| Db | 1561 | AATCCCGCGTCGCCATGGCAACTCACAAGGACGACGAGGACCGCTTTTCCCATCCAG    | 1620 |
| QY | 539  | GlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsn---ThrAlaLeuAspAsn | 557  |
| Db | 1621 | GGAGTCTGTATTTTGGAAAA-----ACTGGAGCACTAACAATACTACATTGGAAAA     | 1674 |
| QY | 558  | ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe    | 577  |
| Db | 1675 | GTGTAATGACAAATGAAGAAGAAATTCGTCCTACTAATCTGTAGCCACGGAAGATA     | 1734 |
| QY | 578  | GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis | 597  |
| Db | 1735 | GGGATAGTCAGCAGCAACTTACAAGCGGCTAATACTGCAGCCCCAGACACAAGTTGT    | 1794 |
| QY | 598  | AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro | 617  |



Dbb 1795 AACcAGGAGcCTTAcCTGGcATGGTCTGGcAGAAcCGGgAcGTGTAcCTGcAGGGTCCc 1854

Qy 618 ILeTrpAlaLySIlleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 637

Dbb 1855 ATCTGGCCAAgATTCTcACAcGGATGGcAACTTcACCCGTCTCTCTTTGATGGCGGC 1914

Qy 638 PheGlyLeuLySAsnProProGlnIlleLeuIlleLySAsnThrProValProAlaAsn 657

Dbb 1915 TTTGGAcTTAAAcATCCGCTCTCTcAGATCCGTGATcAAGAAcAcTCTCCGTCTCCGCTAAcT 1974

Qy 658 ProProAlaGluPheSerAlaThrLySPheAlaSerPheIlleThrGlnTyrSerThrGly 677

Dbb 1975 CCTCCGGAGGTGTTTAcTCTCTCCAAgTTTGCTTCTGTTcATcACAcAGTAcAGcACCGGA 2034

Qy 678 GlnValSerValGluIlleGluTrpGluLeuGlnLySGluAsnSerLySArgTrpAsnPro 697

Dbb 2035 CAAGTCAGCGTGGAAATCGAGTGGGAGCTGCAGAGGAAPAcAGCAAGCGTGGGACCG 2094

Qy 698 GluValGlnTyrThrSerAsnTyrAlaLySSerAlaAsnValAspPheThrValAspAsn 717

Dbb 2095 GAGATTcAGTAcAcCTCCAACTTTGAAAGcAGAcTGGTGTGGAcCTTTGCGGTGAcAGc 2154

Qy 718 AsnGlyLeuTyrThrGluProArgProIlleGlyThrArgTyrLeuThrArgProLeu 736

Dbb 2155 CAGGGTGTTTAcTCTGAGCTCGCCCTATTGGcAcTCTGTTAcCTAcCCCGTAATCTG 2211

RESULT 14

AY243021

LOCUS

DEFINITION Non-human primate Adeno-associated virus isolate AAVch.5 capsid protein (VP1) gene, complete cds.

ACCESSION AY243021

VERSION AY243021.1 GI:29650537

KEYWORDS

SOURCE Non-human primate Adeno-associated virus

ORGANISM Non-human primate Adeno-associated virus

REFERENCE Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.

AUTHORS 1 (bases 1 to 2208)

Gao,G., Alvira,M.R., Somanathan,S., Lu,Y., Vandenbergh,e,L.H., Rux,J.J., Calcedo,R., Sanmiguel,J., Abbas,Z. and Wilson,J.M.

TITLE Adeno-associated viruses undergo substantial evolution in primates during natural infections

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (10), 6081-6086 (2003)

PUBMED 12716974

REFERENCE 2 (bases 1 to 2208)

AUTHORS Gao,G., Alvira,M.R., Somanathan,S., Lu,Y., Vandenbergh,e,L.H., Rux,J., Calcedo,R., Sanmiguel,J., Abbas,Z. and Wilson,J.M.

TITLE Direct Submission

JOURNAL Submitted (23-FEB-2003) Department of Medicine, University of Pennsylvania School of Medicine and Wistar Institute, 204 Wistar Institute, 3601 Spruce Street, Philadelphia, PA 19104, USA

FEATURES

source

1. .2208

/organism="Non-human primate Adeno-associated virus"

/proviral

/mol\_type="genomic DNA"

/isolate="AAVch.5"

/db\_xref="taxon:226582"

1. .2208

/gene="vp1"

1. .2208

/gene="vp1"

/codon\_start=1

/product="capsid protein"

/protein\_id="AAO88207.1"

/db\_xref="GI:29650538"

/translation="MAADGYLPDWLEDTLSEGIQWWKLKPGPPPPKPNQQRHDDSRG

LVLPGYKYLGPENGLDKGEPVNEADAALEHDKAYDHQLKQGNFYLYKNHADAEEFQE

RLOEDTSFGNLRGAVFQAKRVLEPLGLVEEAVKTPAGKKRPIEQSPAEPSSSGIG

KSQQPAPAKRLNFGQGTDESVPDPQPIGEPPAAPSPGVSNWMASSGAGPMADNNEGA

DGVGNSSGNWHCDSTWMDRVITTSRTRWALPTYNNHLYKQISSSEGATPDNDHYFGYS

TPWGYFDNRFHCHFSPRDWQRLLNNNWGFPPKLNFKLFNIQVKEVTQNDGTTTIAN

NLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMIPQYGLTLNNGSQAVGRSSF

YCLEYFPsQMLRTGNFTFSYTFEDVPFHSSYAHsQSdRLMNPliDQYLyLsKTQg

TSgTTQsRLQFSQAGpSSMAQAKNwLPGsYRQORMSKTANDNNSEFAwTAATKY

YLNGRNSLVPgPPMASHKDDeEKYfPMHGNLI fGKQGTGTNNVDIESVLITDEEIR

TTNPVATEQYGVATNHQsQNTTASyGSVDsSQILPGwVWQDRDvYLQGPIwAKTPHT

DGHFHPsPLMGFGGLKHPPpQILIKNTpVPANPATFTTPGKfASfITQYSTGQVSVEI

EWELQKENSKRWNPEIQYTSNYNKsVNVEfTVdANGvYSEPRPIGTRYLTRL"

BASE COUNT 617 a 638 c 518 g 435 t

ORIGIN

Alignment Scores:

Pred. No.: 2.14e-205 Length: 2208

Score: 3423.50 Matches: 619

Percent Similarity: 90.22% Conservative: 45

Best local Similarity: 84.10% Mismatches: 71

Query Match: 85.82% Indels: 1

DB: 14 Gaps: 1

US-09-807-802A-13 (1-736) x AY243021 (1-2208)

Qy 1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20

Dbb 1 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACACTCTCTCTGAAGGAATCAGA 60

Qy 21 GluTrpTrpAspLeuLySProGlyAlaProLySProLySAlaAsnGlnGlnLySGlnAsp 40

Dbb 61 CAGTGGTGAAGCTCAAACCTGGCCcACcACCGCGGAACCTTAACCAACACACCGGGAC 120

Qy 41 AspGlyArgGlyLeuValLeuProGlyTyrLySTyrLeuGlyProPheAsnGlyLeuAsp 60

Dbb 121 GACAGTAGGGGTCTTGTTCTTCTGGGTACAAGTACCTCGGACCCCTTCAACGGACTCGAC 180

Qy 61 LySGlyGluProValAsnAlaAlaAspAlaAlaLeuGluGlyHisAspLySAlaTyrAsp 80

Dbb 181 AAAGGAGAGCCGGTCAACGAGGACAGCCCGCGGCCCTCGAGCACGACAAAGCCTACGAC 240

Qy 81 GlnGlnLeuLySAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100

Dbb 241 CACCAGCTCAAGCAAGGGGACAAcCGGTAcCTCAAAATAcAACcACCGGAcCGTGAATTT 300

Qy 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120

Dbb 301 CAGGAGCGTCTTCAAGAAGATACGTCTTTCGGGGGCACCTCGGGCAGCAGTCTTCCAG 360

Qy 121 AlaLySLySArgValLeuGluProLeuGlyLeuValGluGlyAlaLySThrAlaPro 140

Dbb 361 GCCAAAAAGAGGGTACTCGAGCCCTCTGGTCTGTTGAGGAAGCTGTTAAGACGGCTCCT 420

Qy 141 GlyLySLySArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly 160

Dbb 421 GGAAAAAAGAGACCTATAGAGCAGTCTCTCTGCAGAAcCGGAcTCTTCTCGGGCATCGGC 480

Qy 161 LysThrGlyGlnGlnProAlaLySLySArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180

Dbb 481 AAATCAGGCCAGCAGCCCGCTAAAGAAAGACTCAATTTTGGTCAGACTGGCGACACAGAG 540

Qy 181 SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro 200

Dbb 541 TCAGTCCCAGACCCCTCAACCAATCGGAGAAACCCCGCAGCCCTCTCTGGTGTGGGATCT 600

Qy 201 ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly 220

Dbb 601 AATACAATGGCTTCAGCGGTGGGGcACcCAATGGCAGAcAATAACGAAGCGCGCGACCGA 660

Qy 221 ValGlyAsnAlaSerGlyAsnTyrHisCysAspSerThrTrpLeuGlyAspArgValIle 240

Dbb 661 GTGGGTAACTCTCGGGAAATTGGCATTCGGATTCCACATGGATGGCGACAGAGTTATC 720

Qy 241 ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLySGlnIle 260

Dbb 721 ACCACcAGcACAAGAAcCTGGGcCCCTCCcCACCTAcAATAATcACCTCTAcAAGCAAATC 780

Qy 261 SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp 280

Db 781 TCCAGCGAATCG--GGAGCCACCACGACAACCACTACTTCGGCTACAGCACCCCTGG 837

Qy 281 GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu 300

Db 838 GGGTATTTTGACTTTAACAGATTCCACTGTCTCACTTCTCACCACGTGACTGGCAGCGACTC 897

Qy 301 IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln 320

Db 898 ATCAACAACAACCTGGGATTTAGACCCCAAGAACTCAATTTTCAAGCTCTTCAACATCCAA 957

Qy 321 ValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThr 340

Db 958 GTCAAGGAGGTCAACGAGAATGATGGAAACACGACCATCGCCAAATAACCTTACCAGCACG 1017

Qy 341 valGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln 360

Db 1018 GTGCAGGTCTTCACAGACTCTGAGTACCAGCTGCCCTACGTCCTCGGTCGGCTACCAG 1077

Qy 361 GlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr 380

Db 1078 GGCTGCCTTCGCGGTTCCAGCAGACGCTTTCATGATTCTCTCAGTACGGCTACTTGACT 1137

Qy 381 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 400

Db 1138 CTGAACAATGGCAGCCAAGCGGTAGGACGTTCTTCATTCTACTGTCTAGAGTATTTTCCC 1197

Qy 401 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro 420

Db 1198 TCTCAGATGCTGAGGACGGGAAACAACCTTCACCTTCAGCTACACTTTTGAAGACGTGCCT 1257

Qy 421 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 440

Db 1258 TTCCACAGCAGCTACGCGCACAGCCAGAGTCTGGATCGGCTGATGAATCCTCTCATTTGAC 1317

Qy 441 GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAsp 460

Db 1318 CAGTACCTGTATTACCTTGAGCAAAACTCAGGTACAAAGTGAACAACGACGCAATCGAGA 1377

Qy 461 LeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuPro 480

Db 1378 CTGCAGTTTCAGCCAAGCTGGGCCTAGCTCCATGGCTCAGCAGGCCAAAACCTGGCTACCG 1437

Qy 481 GlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsn 500

Db 1438 GGACCCAGCTACCGACAGCAGCGAATGTCTAAGACGGCTAATGACACACAACAACAGTGAA 1497

Qy 501 PheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnPro 520

Db 1498 TTTGCTTGGACTGCAGCCACCAATATTACCTGAATGGAAGAAATTTCTCGTCAATCCC 1557

Qy 521 GlyThrAlaMetAlaSerHisLysAspAspGluaspLysPhePheProMetSerGlyVal 540

Db 1558 GGGCCCCCAATGGCCAGTCACAAGGACGATGAGAAAGATATTTCCCATGCACGGAAAT 1617

Qy 541 MetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIle 560

Db 1618 CTCATCTTTGGAAAACAAGGCACAGGAACCTACCAATGTGGACATTGAATCAGTGTCTATT 1677

Qy 561 ThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrVal 580

Db 1678 ACAGACGAAGAAGAAATCAGAAACAACCTAATCCTGTGGCTACAGAACATACGGACAGGTT 1737

Qy 581 AlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGly 600

Db 1738 GCCACCAACCATCAGAGTCAGAACACCACACAGCTTCCTATGGAAGTGTGGACAGCCAGGA 1797

Qy 601 AlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAla 620

Db 1798 ATCTTACCTGGAATGGTGTGGCAGGACCGCGATGTCTATCTTCAAGGTCCCATTTGGGCC 1857

Qy 621 LysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeu 640

Db 1858 AAAACTCCTCACACGACGACACTTTTCATCCTCTCCGCTCATGGGAGGCTTTGGACTG 1917

Qy 641 LysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAla 660

Db 1918 AAACACCTCTCCCCAGATCCTGATCAAAAACACACCTGTGCCAGCAATCCCGGACC 1977

Qy 661 GluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSer 680

Db 1978 ACTTTCACCTCTGGAAAGTTGCTTCTGTTCAATACCCAGTATTCCACCGACAGGTGAGC 2037

Qy 681 ValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGln 700

Db 2038 GTGGAATAGAGTGGAGCTGCAGAAAGAAAACAGCAAAACGCTGGAACCCAGAAATTCAG 2097

Qy 701 TyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeu 720

Db 2098 TACACCTCCAACACTACAACAAGTCGGTGAATGTGGAGTTTACCGTGGACGCAACCGTGTT 2157

Qy 721 TyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736

Db 2158 TATTCTGAACCCCGCCCTATTGGCACTCGTTACCTTACCCGGAACCTTG 2205

RESULT 15

AY242998 2190 bp DNA linear VRL 14-MAY-2003

LOCUS Non-human primate Adeno-associated virus isolate AAVrh.37 capsid

DEFINITION protein (VP1) gene, complete cds.

ACCESSION AY242998

VERSION AY242998.1 GI:29650491

KEYWORDS

SOURCE Non-human primate Adeno-associated virus

ORGANISM Non-human primate Adeno-associated virus

REFERENCE 1 (bases 1 to 2190) Gao,G., Alvira,M.R., Somanathan,S., Lu,Y., Vandenberghe,L.H., Rux,J.J., Calcedo,R., Sanmiguel,J., Abbas,Z. and Wilson,J.M. Adeno-associated viruses undergo substantial evolution in primates during natural infections

AUTHORS Proc. Natl. Acad. Sci. U.S.A. 100 (10), 6081-6086 (2003)

TITLE

JOURNAL 12716974

PUBMED 2 (bases 1 to 2190)

REFERENCE Gao,G., Alvira,M.R., Somanathan,S., Lu,Y., Vandenberghe,L.H., Rux,J.J., Calcedo,R., Sanmiguel,J., Abbas,Z. and Wilson,J.M. Direct Submission

AUTHORS Submitted (23-FEB-2003) Department of Medicine, University of Pennsylvania School of Medicine and Wistar Institute, 204 Wistar Institute, 3601 Spruce Street, Philadelphia, PA 19104, USA

TITLE

JOURNAL Location/Qualifiers

FEATURES

source

1..2190 /organism="Non-human primate Adeno-associated virus" /proviral /mol\_type="genomic DNA" /isolate="AAVrh.37" /db\_xref="taxon:226582" 1..2190 /gene="VP1" 1..2190 /gene="VP1" /codon\_start=1 /product="capsid protein" /protein\_id="AAO88184.1" /db\_xref="GI:29650492" /translation="MAADGYLPDWLEDNLSEGIREWWDLKPAGPKPKANQKQDDGRG LVLPGYKYLGPFGNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHDAEFQE RLQEDTSFGGNLGRAVFAQKRVLEPLGLVEEGAKTAPGKRPIDSPDSSTGIGKKGQ QPAKKLNFQGTGDSVDPDQPLGEPAPASPSSVSGTMAAGGAPTADNNEGADGVG NASGNWCHDSITWLGDRVITTSRTWALPTYNNHLYKQISSSSSGATNDNHYFGYSTPW GYFDNRFCHFSPRDWQRLLNNNWGFRPKLRFKLFNIQVKEVTNDGVTTIANNLT STVQVSDSEYQLPYVLSAHQGLCLPPFPADVEMIPOYGYLTLNNGSQSVGRSSFYCL EYFPSQMLRTGNNEFYSFEDVPFHSSYAHQSILDRLMNPLIDQLYLYLARTQSTTG STRELQHQAGPNTWAEQSKNWLPGPCYRQQRSLKNDLFNNNSNFAWTAATKYHLNGR NSLTNPGIPMATNKDDEQFFPINGVLVFGKTGAANKTTLENVLTSEEEKITNPVA TEYGVVSSNLQSTAGPQSQTINSGALPGMWQNRDVLQGPWAKIPIHTDGNFHP SPLMGFGLEHPPPPQLIKNTVPANPPEVFTPAKFASFITQYSTGQVSVEIEWELQK

gene

CDS

ENSKRWNPEIQYTSNYAKSNNVFAVNPDDGVYTEPRPIGTRYLTRNL"

|   |           |  |       |       |
|---|-----------|--|-------|-------|
| BASE COUNT                                    | 558 a     | 662 c  | 545 g | 425 t |
| ORIGIN  |           |  |       |       |
| Alignment Scores:                             |           |  |       |       |
| Pred. No.:                                    | 2.44e-205 | Length:  | 2190  |       |
| Score:  | 3422.50   | Matches:   | 624   |       |
| Percent Similarity:                           | 91.85%    | Conservative:  | 52    |       |
| Best Local Similarity:                        | 84.78%    | Mismatches:  | 53    |       |
| Query Match:                                  | 85.80%    | Indels:  | 7     |       |
| DB:   | 14        | Gaps:  | 3     |       |
| US-09-807-802A-13 (1-736) x AY242998 (1-2190) |           |  |       |       |
| QY  | 1         | MetAlaAlaAspGlyTyrLeuProAspTyrLeuGluAspAsnLeuSerGluGlyIleArg   | 20    |       |
| Db  | 1         | ATGGCTGCCATGGTTATCTTCCAGATTGGCTCGAGGACAAACCTCTCTGAGGCGATTTCGC  | 60    |       |
| QY  | 21        | GluTyrTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp   | 40    |       |
| Db  | 61        | GAGTGGTGGACCTGAAACCTGGAGCCCCGAAACCCAAAGCCACAGCAAAAGCAGGAC      | 120   |       |
| QY  | 41        | AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp   | 60    |       |
| Db  | 121       | GACGGCCGGGTCTGGTCTTCTTGGCTTACAAGTACCTCGGACCTTCAACGGACTCGAC     | 180   |       |
| QY  | 61        | LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp      | 80    |       |
| Db  | 181       | AAGGGGAGCGCGTCAACGGCGGACGAGCGGCCCTCGAGCACGACAAGGCGCTACGAC      | 240   |       |
| QY  | 81        | GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe   | 100   |       |
| Db  | 241       | CAGCAGCTCAAAGCGGTGACAATCCGTACTCGGTATTAACACGCGACGCCGAGTTT       | 300   |       |
| QY  | 101       | GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln   | 120   |       |
| Db  | 301       | CAGGAGCGCTGCAAGAAGATACGTCATTTGGGGCAACCTCGGGCGAGCAGTCTTCCAG     | 360   |       |
| QY  | 121       | AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro      | 140   |       |
| Db  | 361       | GCCAAAGAAGCGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGGCGCTAAGACGGTCTCT   | 420   |       |
| QY  | 141       | GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly   | 160   |       |
| Db  | 421       | GGAAGAAGAGACCCATAGAC-----TCTCCAGACTCTCCACGGGCATCGGC            | 468   |       |
| QY  | 161       | LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu   | 180   |       |
| Db  | 469       | AAAAAAGGCCAGCAGCCCGCTAAAAAAGAAGCTCAATTTGGTCAGACTGGCGACTCAGAG   | 528   |       |
| QY  | 181       | SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro   | 200   |       |
| Db  | 529       | TCAGTCCCCGACCCCTCAACCTCTTGGAGAACCTCCAGCAGCGCCCTCTAGTGGGATCT    | 588   |       |
| QY  | 201       | ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGlnGlyAlaAspGly      | 220   |       |
| Db  | 589       | GGTACAAATGGCTCGAGCGGTGGCGCACCAACGCGCAGACAATAACGAAGGTGCCGACGGA  | 648   |       |
| QY  | 221       | ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle   | 240   |       |
| Db  | 649       | GTGGGTAATGCCCTCAGGAATTTGGCATTTCCATGCTGGCTGGCGCAGAGATCATC       | 708   |       |
| QY  | 241       | ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle   | 260   |       |
| Db  | 709       | ACCACCAGCACCAACCTGGGCCCTCCCCACCTACAACAAACCACCTCTACAAGCAAAATC   | 768   |       |
| QY  | 261       | SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp   | 280   |       |
| Db  | 769       | TCCAGCAGCAGCTCAGGAGCCACCAATGACAACCACTACTTCGGCTACAGCACCCTCTGG   | 828   |       |
| QY  | 281       | GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTyrGlnArgLeu   | 300   |       |
| Db  | 829       | GGGTATTTTGACTTTTAACAGATTCCACTGCCACTTCTCACCACGTTGACTGGCAGCGACTC | 888   |       |

|    |      |  |      |  |
|----|------|--|------|--|
| QY | 301  | IleAsnAsnAsnTyrGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln   | 320  |  |
| Db | 889  | ATCAACAACAACACTGGGATTCGGGCCCAAGAAGCTCGGGTTCAAGCTCTTCAACATCCAG  | 948  |  |
| QY | 321  | VallysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThr   | 340  |  |
| Db | 949  | GTCAAGGAGGTCAACAACGATGACGGCGTCAAGACCATCGCTAATAACCTTACCAGCACG   | 1008 |  |
| QY | 341  | ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln   | 360  |  |
| Db | 1009 | GTTCAGGTCTTCTCGGACTTCGGAATACCAGCTGCCGTACGTCTCGGCTCTGCGCACCAG   | 1068 |  |
| QY | 361  | GlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr   | 380  |  |
| Db | 1069 | GGTGCCTGCCTCCGTTCCCGCGGACGCTTTCATGATTCCTCAGTACGGCTACCTGACT     | 1128 |  |
| QY | 381  | LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro   | 400  |  |
| Db | 1129 | CTGAACAACGGCAGCCAATCGGTGGCGCGTTCTCTCTTCTACTGCCTGGAATATTTCCCC   | 1188 |  |
| QY | 401  | SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro   | 420  |  |
| Db | 1189 | TCTCAAATGCTGAGAACGGGCAACAACCTTTGAGTTTCAGTTACAGCTTCGAGGACGTGCCT | 1248 |  |
| QY | 421  | PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp   | 440  |  |
| Db | 1249 | TTCCACAGCAGCTACGCGCACAGCCAGAGCCTAGACCGGCTGATGAACCTCTCATCGAC    | 1308 |  |
| QY | 441  | GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAsp   | 460  |  |
| Db | 1309 | CAGTACCTGTACTACCTGGCCCGGACCCAGAGCACCCACGGGTTCC-----ACCAGGGA    | 1362 |  |
| QY | 461  | LeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuPro   | 480  |  |
| Db | 1363 | CTGCAATTCATCAAGCTGGGCCCAATACTATGGCCGAGCAGTCAAGAAGACTGGTGCCT    | 1422 |  |
| QY | 481  | GlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnSerAsn      | 500  |  |
| Db | 1423 | GGACCTGCTATAGGCAACAGAGACTGTCAAAGAACTTGGACTTTAAACAACAACAGCAAT   | 1482 |  |
| QY | 501  | PheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnPro   | 520  |  |
| Db | 1483 | TTTGCCTGGACTGCTGCCACTAAATATCATCTGAATGGCAGAAACTTTTGACCAATCCT    | 1542 |  |
| QY | 521  | GlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePhePrometSerGlyVal   | 540  |  |
| Db | 1543 | GGCATTTCCCATGGCAACCAACAGGATGATGAGGACCAGTCTTTTCCCATCAACGGGGTA   | 1602 |  |
| QY | 541  | MetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIle   | 560  |  |
| Db | 1603 | CTGGTTTGGCAAGACGGGAGCT---GCCAACAAACTACGCTGGAAACGTTCTGATG       | 1659 |  |
| QY | 561  | ThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrVal   | 580  |  |
| Db | 1660 | ACCAGCAGGAGGAGATCAAGACCCTAACCTGTGGCTACAGAAGAATACGGTGTGGTC      | 1719 |  |
| QY | 581  | AlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGly   | 600  |  |
| Db | 1720 | TCCAGCAACCTGCAGTCGTCTACAGCGGGGCTCAATTCACAGACTATCAACAGCCAGGGA   | 1779 |  |
| QY | 601  | AlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAla   | 620  |  |
| Db | 1780 | GCATGCTGCGCATGGTCTGGCAGAACCGGACGCTGTATCTGAGGCTCCCATCTGGGCC     | 1839 |  |
| QY | 621  | LysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyPheGlyLeu      | 640  |  |
| Db | 1840 | AAAATTCCTCACACGGATGGCAACTTTCACCCGCTCTCCTCTGATGGCGGTTTGGACTC    | 1899 |  |
| QY | 641  | LysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAla      | 660  |  |
| Db | 1900 | GAACACCCCGCTCCACAGATCCTGTATCAAAAACACACCTGTACTCTGCTAATCTCCGGAG  | 1959 |  |



|    |      |   |      |
|----|------|---|------|
| QY | 661  | GlupheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSer  | 680  |
| Db | 1960 | GTGTTTACTCCTGCCAAGTTTGCCCTCCTTCATCAGCGAGTACAGCACCGGACAAGTCAGC | 2019 |
| QY | 681  | ValGluIleGluTyrGluLeuGlnLysGluAsnSerLysArgTyrAsnProGluValGln  | 700  |
| Db | 2020 | GTGGAATCGAGTGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAACCCAGAAATTCAG   | 2079 |
| QY | 701  | TyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeu  | 720  |
| Db | 2080 | TATACCTCCAATTATGCCAAGTCTAATAATGTTGAATTGCTGTGAACCCCTGATGGTGTT  | 2139 |
| QY | 721  | TyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu              | 736  |
| Db | 2140 | TATACTGAGCCTCGCCCCCATTTGGCACTCGTTACCTCACCCCGTAATCTG           | 2187 |

Search completed: February 14, 2004, 22:37:19  
Job time : 7686.62 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 14, 2004, 15:18:00 ; Search time 504.843 Seconds  
(without alignments)  
3935.452 Million cell updates/sec

Title: US-09-807-802A-13  
Perfect score: 3989  
Sequence: 1 MAADGYLPDWLEDNLSGIR.....NNGLYTEPRPIGTRYLTRPL 736

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -MODEL=frame+\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09807802/runat\_11022004\_175606\_15919/app.query.fasta\_1.2389  
-DB=N Geneseq\_19Jun03 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09807802@cgn 1 1 806 @runat\_11022004\_175606\_15919 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq\_19Jun03: \*  
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: \*  
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: \*  
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: \*  
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: \*  
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: \*  
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: \*  
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: \*  
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: \*  
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: \*  
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: \*  
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: \*  
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: \*  
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: \*  
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: \*  
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: \*  
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: \*  
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: \*  
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: \*  
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: \*  
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: \*  
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: \*  
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: \*  
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: \*  
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: \*  
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Match | Length | DB | ID       | Description        |
|------------|--------|-------|--------|----|----------|--------------------|
| 1          | 3989   | 100.0 | 2211   | 21 | AAD00777 | Adeno-associated v |
| 2          | 3989   | 100.0 | 4718   | 21 | AAD00772 | Adeno-associated v |
| 3          | 3963   | 99.3  | 4683   | 22 | AAF23749 | AAV6 DNA sequence. |
| 4          | 3511   | 88.0  | 4722   | 22 | AAF23748 | AAV3B DNA sequence |
| 5          | 3439   | 86.2  | 8178   | 21 | AAD00834 | Chimeric adeno-ass |
| 6          | 3402.5 | 85.3  | 4072   | 22 | AAH26327 | Adenovirus helper  |
| 7          | 3402.5 | 85.3  | 4679   | 22 | AAI66974 | Adeno-associated v |
| 8          | 3402.5 | 85.3  | 4679   | 22 | AAF23750 | AAV2 DNA sequence. |
| 9          | 3402.5 | 85.3  | 4679   | 24 | ABK89694 | Adeno-associated v |
| 10         | 3402.5 | 85.3  | 4679   | 25 | ABV76133 | Adeno-associated v |
| 11         | 3402.5 | 85.3  | 7557   | 22 | AAH26326 | Adeno-associated v |
| 12         | 3402.5 | 85.3  | 8698   | 22 | AAH26325 | Wild-type adeno-as |
| 13         | 3360.5 | 84.2  | 4680   | 17 | AAT09008 | Wild-type adeno-as |
| 14         | 3360.5 | 84.2  | 4680   | 25 | ABX11497 | Wild-type adeno-as |
| 15         | 3322.5 | 83.3  | 4675   | 22 | AAH41481 | Adeno-associated v |
| 16         | 3322.5 | 83.3  | 4675   | 22 | AAF89931 | Nucleotide sequenc |
| 17         | 3322.5 | 83.3  | 4675   | 24 | ABS69879 | Human adeno-associ |
| 18         | 3322.5 | 83.3  | 4675   | 24 | ABS69880 | Human adeno-associ |
| 19         | 3322.5 | 83.3  | 4675   | 24 | ABA02989 | Adeno-associated v |
| 20         | 3251   | 81.5  | 1800   | 21 | AAD00778 | Adeno-associated v |
| 21         | 3038.5 | 76.2  | 8151   | 21 | AAD00832 | Chimeric adeno-ass |
| 22         | 2906   | 72.9  | 1605   | 21 | AAD00779 | Adeno-associated v |
| 23         | 2486.5 | 62.3  | 2208   | 19 | AAV21650 | AAV4 VP1 capsid pr |
| 24         | 2486.5 | 62.3  | 2208   | 25 | ABX15809 | Adeno-associated v |
| 25         | 2481.5 | 62.2  | 4767   | 19 | AAV21648 | AAV4 genome. Aden  |
| 26         | 2481.5 | 62.2  | 4767   | 25 | ABX15807 | Adeno-associated v |
| 27         | 2476.5 | 62.1  | 7215   | 21 | AAD00831 | Hybrid adeno-assoc |
| 28         | 2220   | 55.7  | 2307   | 21 | AAZ49212 | Adeno associated v |
| 29         | 2220   | 55.7  | 4652   | 21 | AAZ49209 | Adeno associated v |
| 30         | 2220   | 55.7  | 4652   | 24 | AAS17712 | Adeno-associated v |
| 31         | 2214   | 55.5  | 2264   | 21 | AAZ49213 | Adeno-associated v |
| 32         | 2214   | 55.5  | 2264   | 21 | AAZ49214 | Adeno associated v |
| 33         | 2195.5 | 55.0  | 2487   | 18 | AAT69854 | Duck parvovirus ca |
| 34         | 2168.5 | 54.4  | 4570   | 16 | AAT01190 | Barbary duck parvo |
| 35         | 1830.5 | 45.9  | 1800   | 19 | AAV21657 | AAV4 VP2 coat prot |
| 36         | 1830.5 | 45.9  | 1800   | 25 | ABX15816 | Adeno-associated v |
| 37         | 1690.5 | 42.4  | 1617   | 19 | AAV21658 | AAV4 VP3 coat prot |
| 38         | 1690.5 | 42.4  | 1617   | 25 | ABX15817 | Adeno-associated v |
| 39         | 1410   | 35.3  | 2271   | 21 | AAD00833 | Adeno-associated v |
| 40         | 921    | 23.1  | 606    | 24 | ABK33262 | Adeno-associated v |
| 41         | 606.5  | 15.2  | 2343   | 20 | AAH81583 | Erythrovirus V9 DN |
| 42         | 606.5  | 15.2  | 5028   | 20 | AAH81580 | Genomic DNA sequen |
| 43         | 605.5  | 15.2  | 4677   | 16 | AAT49535 | Human parvovirus 9 |
| 44         | 599.5  | 15.0  | 2380   | 25 | ABZ59573 | Human parvovirus B |
| 45         | 599.5  | 15.0  | 2380   | 25 | ABZ59576 | Human parvovirus B |

ALIGNMENTS

RESULT 1  
AAD00777  
ID AAD00777 standard; DNA; 2211 BP.  
XX  
AC AAD00777;  
XX  
DT 08-SEP-2000 (first entry)  
XX  
DE Adeno-associated virus serotype 1 capsid protein VP1 DNA.  
XX  
KW Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;  
KW cap protein; recombinant viral vector; gene delivery; gene therapy;  
KW vaccine; transgene; VP1; ss.  
XX  
OS Adeno associated virus serotype 1.  
XX  
FH Key Location/Qualifiers

```
CDS      1..2211
FT       /*tag= a
FT       /product= "VP1 protein"
XX
XX WO200028061-A2.
XX
XX 18-MAY-2000.
XX
XX 02-NOV-1999; 99WO-US25694.
XX
XX 05-NOV-1998; 98US-0107114.
XX
XX (UYPE-) UNIV PENNSYLVANIA.
XX
XX Wilson JM, Xiao W;
XX
XX WPI; 2000-376571/32.
XX
XX P-PSDB; AAY71167.
XX
XX Novel adeno-associated virus serotype 1 polynucleotide useful for
XX preparation of medicament for delivery of a transgene to a host -
XX
XX Claim 10; Page 83-87; 108pp; English.
XX
XX The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA
XX which is characterised by two inverted terminal repeats (ITR) and open
XX reading frames for rep and capsid (cap) proteins. The rep reading frame
XX encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap
XX reading frame encodes three structural proteins, VP1, VP2 and VP3.
XX The AAV-1 sequence or its fragments particularly ITRs, rep and cap
XX coding regions, are useful in production of recombinant viral vectors
XX for gene delivery. These vectors can be used as gene therapy
XX vectors, vaccine vectors or antisense delivery vectors. The AAV-1 does
XX not induce the formation of neutralising antibodies specific to any
XX serotype of AAV hence is useful for transducing host cells, and in the
XX preparation of a medicament for the delivery of transgene to a host.
XX The present sequence is an AAV-1 DNA encoding a cap protein VP1 which is
XX useful in the production of recombinant viral vector for gene delivery.
XX
XX Sequence 2211 BP; 577 A; 652 C; 538 G; 444 T; 0 other;
SQ

Alignment Scores:
Pred. No.:      2.37e-283      Length:      2211
Score:          3989.00        Matches:      736
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:     100.00%      Indels:        0
DB:              21           Gaps:         0

US-09-807-802A-13 (1-736) x AAD00777 (1-2211)

QY      1 MetAlaAlaAspGlyTyrLeuProAspTyrLeuGluAspAsnLeuSerGluGlyIleArg 20
Db      1 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTCGC 60
QY      21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp 40
Db      61 GAGTGGTGGGACTTGAAACCTGGAGCCCCGAGCCCAAGCCCAACGAGCAAAAGCAGGAC 120
QY      41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp 60
Db      121 GACGGCCGGGCTCTGGTGCTCTCTGGCTACAGTACCTCGGACCTTCAACGGACTCGAC 180
QY      61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80
Db      181 AAGGGGAGCCCGTCAACGGCGGCGAGCGCGCCCTCGAGCAGCAGCAAGGCTACGAC 240
QY      81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100
Db      241 CAGCAGCTCAAAGCGGGTGACAAATCCGTACCTCGGTATTAACCAACGCCCGCGCGAGTTT 300
QY      101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120
Db      1381 TTGCTGTTAGCCCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCCAAAACACTGGCTACCT 1440
```

```
Db      301 CAGGAGCGTCTGCAAGAAGATACGTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG 360
QY      121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro 140
Db      361 GCCAAGAAGCGGGTCTCGAAACCTCTCGGTCTGGTTGAGGAAGGCGCTAAGACGGCTCCT 420
QY      141 GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly 160
Db      421 GGAAGAAGAACGTCGCGGTAGAGCAGTCGCCACAAGAGCAGACTCTCTCTCGGGCATCGGC 480
QY      161 LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180
Db      481 AAGACAGGCCAGCAGCCCGCTAAAAAAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG 540
QY      181 SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro 200
Db      541 TCAGTCCCCGATCCACAACTCTCGGAGAACCTCCAGCAACCCCGCTGCTGTGGGACCT 600
QY      201 ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly 220
Db      601 ACTACAATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGCGCGCGACGGA 660
QY      221 ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle 240
Db      661 GTGGTAATGCCTCAGGAATTTGGCATTGGCATTCGATTCATGCTGGCGACAGAGTCATC 720
QY      241 ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle 260
Db      721 ACCACGACACCCGCCACCTGGGCTTGGCCACCTACATAACCACTCTTACAAGCAATC 780
QY      261 SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp 280
Db      781 TCCAGTGTCTTCAACGGGGGCGAGCAACGACAACCACTACTTCGGCTACAGCACCCCTGG 840
QY      281 GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu 300
Db      841 GGGTATTTGATTTCAACAGATTCCACTGCCACTTTTCAACCACTGACTGGCAGCGACTC 900
QY      301 IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln 320
Db      901 ATCAACAACAATTTGGGATTCGCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 960
QY      321 ValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThr 340
Db      961 GTCAAGGAGGTCAACGACGAATGATGGCGTCACAACCATCGTCTAATAACCTTACCAGCAG 1020
QY      341 ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln 360
Db      1021 GTTCAAGTCTTCTCGGACTCGGAGTACCAGCTTCCGTACGTTCTCTGGCTCTGCGCACCAG 1080
QY      361 GlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr 380
Db      1081 GGCTGCCTCCCTCCGTTCCCGGGGACGCTGTTCATGATTTCCGCAATACGGCTACCTGACG 1140
QY      381 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 400
Db      1141 CTCAACAATGGCAGCCCAAGCCGTGGGACGTTTCATCTCTTTTACTGCTGGAATATTTCCCT 1200
QY      401 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro 420
Db      1201 TCTCAGATGCTGAGAACGGGCAACAACCTTACCTTCAGCTACACCTTTGAGGAAGTGCCT 1260
QY      421 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 440
Db      1261 TTCCACAGCAGCTACGCGCACAGCAGAGCCCTGGACCGCTGATGAATCCTCTCATCGAC 1320
QY      441 GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAsp 460
Db      1321 CAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCCGGGAAGTGCCCAAAACAAGGAC 1380
QY      461 LeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuPro 480
Db      1381 TTGCTGTTAGCCCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCCAAAACACTGGCTACCT 1440
```





```
FT /product= "VP3 protein"
FT /note= "Capsid protein"
FT 4447..4452
FT polyA_signal
FT /*tag= u
FT 4576..4718
FT /*tag= v
FT /label= 3'_ITR
FT /note= "Inverted terminal repeat which is capable
FT of forming T-shaped hairpin structure"
XX
PN WO200028061-A2.
XX
XX 18-MAY-2000.
PD
XX
PF 02-NOV-1999; 99WO-US25694.
XX
XX 05-NOV-1998; 98US-0107114.
PR
XX (UYPE-) UNIV PENNSYLVANIA.
PA
XX Wilson JM, Xiao W;
PI
XX WPI; 2000-376571/32.
DR
DR P-PSDB; AAY71161, AAY71164, AAY71165, AAY71166, AAY71167, AAY71168,
DR AAY71169.
XX
PT Novel adeno-associated virus serotype 1 polynucleotide useful for
PT preparation of medicament for delivery of a transgene to a host -
XX
PS Claim 1; Fig 1; 108pp; English.
XX
CC The present sequence is an adeno-associated virus serotype 1 (AAV-1)
CC DNA characterised by two inverted terminal repeats (ITR) and open
CC reading frames for rep and capsid (cap) proteins. The rep reading frame
CC encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap
CC reading frame encodes three structural proteins, VP1, VP2 and VP3.
CC The AAV-1 sequence or its fragments particularly ITRs, rep and cap
CC coding regions, are useful in production of recombinant viral vectors
CC for gene delivery. These vectors can be used as gene therapy
CC vectors, vaccine vectors or antisense delivery vectors. The AAV-1 does
CC not induce the formation of neutralising antibodies specific to any
CC serotype of AAV hence is useful for transforming host cells, and in the
CC preparation of a medicament for the delivery of transgene to a host.
XX
SQ Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 other;

Alignment Scores:
Pred. No.: 6.23e-283 Length: 4718
Score: 3989.00 Matches: 736
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

US-09-807-802A-13 (1-736) x AAD00772 (1-4718)

QY 1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20
Db 2223 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTCGC 2282
QY 21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnLysGlnAsp 40
Db 2283 GAGTGGTGGGACTTGAAACCTGGAGCCCCGAAGCCCCAAAGCCCAACCAGCAAAAGCAGGAC 2342
QY 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp 60
Db 2343 GACGGCCGGGTTGGTGCTTCTCTGGCTACAAGTACCTCGGACCCCTTCAACGGACTCGAC 2402
QY 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80
Db 2403 AAGGGGGAGCCCGTCAACGGCGGACGCGCGGCCCTCGAGCAGCAAGGCCTACGAC 2462
QY 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100
```

```
Db 2463 CAGCAGCTCAAAGCGGGTGACAAATCCGTACCTGGGTATAACCAAGCCGACGCCGAGTTT 2522
QY 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120
Db 2523 CAGGAGCGTCTGCAGAAGAATACGTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG 2582
QY 121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro 140
Db 2583 GCCAAGAAGCGGGTTCTCGAAACCTCTCGGTCTGGTGAAGGAGCGCTAAGACGGCTCCT 2642
QY 141 GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly 160
Db 2643 GGAAGAAGAACGTCCGGTAGAGCAGTCGCCACAAAGAGCCAGACTCCTCTCGGGCATCGGC 2702
QY 161 LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180
Db 2703 AAGACAGGCCAGCAGCCCGCTAAAAAGAGACTCAATTTGGTTCAGACTGGCGACTCAGAG 2762
QY 181 SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro 200
Db 2763 TCAGTCCCGATCCACAACCTCTCGGAGAACCTCCAGCAACCCCGCTGCTGTGGGACCT 2822
QY 201 ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly 220
Db 2823 ACTACAATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCCCGCAGGA 2882
QY 221 ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle 240
Db 2883 GTGGGTAATGCCTCAGGAAATTTGGCATTTGCCGATTTCACATGGCTGGCGCAGAGTCATC 2942
QY 241 ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle 260
Db 2943 ACCACCAGCACCCGACCTGGGCTTGGCCACCTTACAATAACCACTCTACAAGCAATC 3002
QY 261 SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp 280
Db 3003 TCCAGTGCTTCAACGGGGCCAGCAACGACAACCACTACTTCGGCTACAGCACCCCTGG 3062
QY 281 GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu 300
Db 3063 GGGTATTTTGATTTCAACAGATTCCACTGCCACTTTTCAACACGCTGACTGGCAGCGACTC 3122
QY 301 IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln 320
Db 3123 ATCAACAACAATTGGGGATTCCGGCCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 3182
QY 321 ValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThr 340
Db 3183 GTCAAGGAGGTACGACGAATGATGGCGTCACAACCATCGCTAATAACCTTACCAGCAGC 3242
QY 341 ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln 360
Db 3243 GTTCAAGTCTTCTCGGACTCGGAGTACCAGCTTCCGTACGTCTCGGCTCTCGGCACCAG 3302
QY 361 GlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr 380
Db 3303 GSGCTGCCTCCCTCCGTTCCGGCGGACGCTGTTTCATGATTCCGCAATACGGCTACCTGACG 3362
QY 381 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 400
Db 3363 CTCAACAATGGCAGCCCAAGCGTGGGACGTTTCATCTCTTTACTGCTGGAATATTTCCCT 3422
QY 401 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro 420
Db 3423 TCTCAGATGCTGAGAACGGGCAACAACCTTTACCTTCAGTACACCTTTGAGGAAGTGCCT 3482
QY 421 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 440
Db 3483 TTCACAGCAGCTACGGGCACAGCCAGAGCCTGGACCGGCTGATGAATCTCTCATCGAC 3542
QY 441 GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAsp 460
```

Db 3543 CAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTCCCCAAACAAGGAC 3602

Qy 461 LeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuPro 480

Db 3603 TTGCTGTTAGCGGTGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAACTGGTACCT 3662

Qy 481 GlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnSerAsn 500

Db 3663 GGACCCCTGTATCGGCAGCAGCGCGTTTCTAANAACAAAACAGACAACAACAGCAAT 3722

Qy 501 PheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleAsnPro 520

Db 3723 TTTACCTGGACTGGTGTCTTCAAAATATAAACCTCAATGGCGTGAATCCATCAACCT 3782

Qy 521 GlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyVal 540

Db 3783 GGCACCTGCTATGGCCTCACACAAAGACGACGAAGACAAGTCTCTTTCCCATGAGCGGTGC 3842

Qy 541 MetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIle 560

Db 3843 ATGATTTTTGGAAAAAGAGAGCGCGGAGCTTCAACACTGCATTGGACAATGTCATGATT 3902

Qy 561 ThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrVal 580

Db 3903 ACAGACGAAGAGGAATTAAGCCACTAACCTGTGGCCACCAGAAAGATTTGGACCGTG 3962

Qy 581 AlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGly 600

Db 3963 GCAGTCAATTTCCAGAGCAGCAGCACAGACCCTGCGACCGGAGATGTCATGCTATGGGA 4022

Qy 601 AlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAla 620

Db 4023 GCATTACCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCAGGTCCTCATTTGGGCC 4082

Qy 621 LysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeu 640

Db 4083 AAAATTCTCACACAGATGGACACTTTCACCCGTCTCTCTTATGGCGGCTTTGGACTC 4142

Qy 641 LysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAla 660

Db 4143 AAGAACCCGCTCTCTCAGATCCTCATCAAAAACACGCTGTCTCTCGGAATCCTCCGGCG 4202

Qy 661 GluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSer 680

Db 4203 GAGTTTTCAGCTACAAAGTTTGCTTCAATTCATACCCCAATACTCCACAGGACAAGTGAGT 4262

Qy 681 ValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGln 700

Db 4263 GTGGAATTTGAATGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAATCCCGAAGTGCAG 4322

Qy 701 TyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeu 720

Db 4323 TACACATCCAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAACAATGGACTT 4382

Qy 721 TyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736

Db 4383 TATACTGAGCCTCGCCCATTTGGCACCCCGTTACCTTACCCGTCCCTG 4430

RESULT 3

AAF23749

ID AAF23749 standard; DNA; 4683 BP.

XX AAF23749;

AC AAF23749;

XX 28-MAR-2001 (first entry)

XX AAV6 DNA sequence.

XX AAV6; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;

KW athrosclerosis; sickle cell anaemia; thalassaemia;

KW blood clotting disorder; diabetes; ss.

XX Adeno associated virus.

XX US6156303-A.

PN 05-DEC-2000.

XX 11-JUN-1997; 97US-0873168.

PF 11-JUN-1997; 97US-0873168.

XX (UNIW ) UNIV WASHINGTON.

PA Russell DW, Rutledge EA;

PI WPI; 2001-060164/07.

XX Adeno-associated virus serotype 6 and viral vector derived from it for gene therapy of cystic fibrosis, cancer, acquired immunodeficiency syndrome, sickle cell anemia, thalassemia and diabetes

PS Claim 1; Fig 1; 50pp; English.

XX The present invention relates to adeno-associated virus serotypes. The present sequence is the DNA sequence of one such serotype (AAV6). AAV6 can be used to construct AAV viral vectors for use in gene therapy for a range of disorders: cystic fibrosis, cancer, AIDS, athrosclerosis, sickle cell anaemia, thalassaemia, blood clotting disorders and diabetes. The AAV viral vectors have increased transduction efficiency of a particular host cell as the AAV virion containing the AAV vector genome can be modified to express a capsid protein of an AAV serotype that transduces the selected host cell.

XX Sequence 4683 BP; 1114 A; 1363 C; 1277 G; 929 T; 0 other;

SQ

Alignment Scores:

Pred. No.: 5e-281 Length: 4683

Score: 3963.00 Matches: 730

Percent Similarity: 99.59% Conservative: 3

Best Local Similarity: 99.18% Mismatches: 3

Query Match: 99.35% Indels: 0

DB: 22 Gaps: 0

US-09-807-802A-13 (1-736) x AAF23749 (1-4683)

Qy 1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20

Db 2208 ATGGTGGCGGATGGTTATCTTCAGATTGGTCTCGAGGACAACTCTCTGAGGGCATTCGC 2267

Qy 21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnLysGlnAsp 40

Db 2268 GAGTGTGGGACTTGAAACCTGAGGCCCGAAACCCAAAGCAACCAAGCAGGAC 2327

Qy 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyPropheAsnGlyLeuAsp 60

Db 2328 GACGGCGGGGTCTGGTGTCTTCTGGTACAAGTACCTCGGACCTTCAACGGACTCGAC 2387

Qy 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80

Db 2388 AAGGGGAGCCCCGTCAACGGCGGGATGCAGCGGCCCTCGAGCAGCAAGGCTACGAC 2447

Qy 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100

Db 2448 CAGCAGCTCAAAAGCGGGTGACATCCGTACCTCGGTATACCCACGCCGCGAGTTT 2507

Qy 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120

Db 2508 CAGGAGCGTCTGCAAGAAGATACGTCTTTTGGGGCAACCTCGGGCGAGCAGTCTTCCAG 2567

Qy 121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGluGlyAlaLysThrAlaPro 140

Db 2568 GCCAAGAAGAGGGTTCTCGAACCTTTTGGTCTGGTTGAGGAAGGTGCTAAGACGGCTCCT 2627

Qy 141 GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerGlyIleGly 160



Db 2628 GGAAAGAAACGTCGGTAGAGCAGTCGCCACAAGAGCCAGACTCCTCTCGGGCATTTGGC 2687

Qy 161 LysThrGlyGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180

Db 2688 AAGACAGGCCAGCAGCCCGCTAAAAAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG 2747

Qy 181 SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro 200

Db 2748 TCAGTCCCGACCCACAACCTCTCGAGAACCTCCAGCAACCCCGCTGCTGTGGGACCT 2807

Qy 201 ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly 220

Db 2808 ACTACAATGGCTTCAGCGGTGGCGCAACCTTCCAGCAACAATAACGAAGCGCCGACCGA 2867

Qy 221 ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValile 240

Db 2868 GTGGGTAAATGGCTCAGGAATTTGGCAATTCGATTCCACATGGCTGGCGCAGAGTCATC 2927

Qy 241 ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnile 260

Db 2928 ACCACCAGCACCCGAACATGGGCCCTTGCCACCCTATAACAACCACTCTACAAGCAAAATC 2987

Qy 261 SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp 280

Db 2988 TCCAGTGTCTCAACGGGGCCAGCAACGACAACCACTACTTCGCTACAGCACCCCTGG 3047

Qy 281 GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu 300

Db 3048 GGGTATTTTGATTTCAACAGATTCCACTGCCATTCTCACCCAGTGACTGGCAGCGACTC 3107

Qy 301 IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln 320

Db 3108 ATCAACAACAATTTGGGATTCCGGCCCCAAGAGACTCAACTTCAAGCTCTTCAACATCCAA 3167

Qy 321 ValLysGluValThrThrAsnAspGlyValThrIleAlaAsnAsnLeuThrSerThr 340

Db 3168 GTCAAGGAGGTACGACGAATGATGGCGTCAGACCATCGTCTAATAAACCTTACCAGCACG 3227

Qy 341 ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln 360

Db 3228 GTTCAAGTCTTCTCGGACTCGGAGTACCAGTTGCCGTACGTCCTCGGCTCTGCGCACCG 3287

Qy 361 GlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr 380

Db 3288 GGCTGCCTCCCTCCGTTCCCGGCGGACGTGTTCATGATTCCGCAGTACGGCTACCTAACG 3347

Qy 381 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 400

Db 3348 CTCACAATGGCAGCCAGGCAGTGGGACGGTCATCCTTTTACTGCTGGAATATTTCCCA 3407

Qy 401 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro 420

Db 3408 TCGCAGATGCTGAGAACGGGCAATAACTTTACCTTCAGCTACACCTTCGAGGACGTGCCT 3467

Qy 421 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 440

Db 3468 TTCCACAGCAGCTACGCGCACAGCCAGAGCCTGGACCGGTGATGAATCCTCTCATCGAC 3527

Qy 441 GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAsp 460

Db 3528 CAGTACCTGTATTACCTGAACAGAACTCAGAATCAGTCCGGAAGTGCCCAAAAACAAGGAC 3587

Qy 461 LeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuPro 480

Db 3588 TTGCTGTTTAGCCGGGGTCTCCAGCTGGCATGTCTGTTACGCCCAAAAACGTGGCTACCT 3647

Qy 481 GlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsn 500

Db 3648 GGACCCTGTTACCGGCGCAGCGCGGTTTCTAAACAANAACAGACAACAACAACAGCAAC 3707

Qy 501 PheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnPro 520

Db 3708 TTTACCTGGACTGGTGTCTCAAAATATAACTTAAATGGGCGTGAATCTATAATCAACCT 3767

Qy 521 GlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyVal 540

Db 3768 GGCACCTGCTATGGCTCACACAAAGACGACAAAGACAAGTTCTTTCCCATGAGCGGTGTC 3827

Qy 541 MetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIle 560

Db 3828 ATGATTTTGGAAAGGAGAGCGCCGGAGCTTCAAACACTGCATTTGGACAATGTCATGATC 3887

Qy 561 ThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrVal 580

Db 3888 ACAGACGAAGAGGAATCAAAGCCACTAACCCCGTGGCCACCGAAAGATTTGGGACTGTG 3947

Qy 581 AlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGly 600

Db 3948 GCAGTCAATCTCCAGAGCAGCAGCACAGACCCTCGACCGGAGATGTGCATGTTATGGGA 4007

Qy 601 AlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAla 620

Db 4008 GCCTTACCTGGAATGGTGTGGCAAGACAGAGACGTATACCTGCAGGTCCTATTGGGCC 4067

Qy 621 LysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeu 640

Db 4068 AAAATTCCTCACACGGATGGACACTTTTCAACCCGCTCTCTCTCATGGGGCGCTTGGACTT 4127

Qy 641 LysAsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAla 660

Db 4128 AAGCACCCGCTCCTCAGATCCTCATCAAAAACACGCTGTTCTCGGAATCCTCCGGCA 4187

Qy 661 GluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSer 680

Db 4188 GAGTTTTCGGCTACAAAGTTTGCTTTCATTTCATCACCCAGTATCCACAGGACAAGTGAGC 4247

Qy 681 ValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGln 700

Db 4248 GTGGAGATTGAATGGGAGCTGCAGAAAGAAAACAGCAAAACGCTGGAATCCCGAAGTGAC 4307

Qy 701 TyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeu 720

Db 4308 TATACATCTAATATGCAAAATCTGCCAACGTTGATTTCACCTGTGGACAACAATGGACTT 4367

Qy 721 TyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736

Db 4368 TATACTGAGCTCGCCCCCATTTGGCACCCGTTACCTCACCCGTCCTCCCTG 4415

RESULT 4

AAF23748

ID AAF23748 standard; DNA; 4722 BP.

XX AAF23748;

XX 28-MAR-2001 (first entry)

XX AAV3B DNA sequence.

DE AAV3B; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;

XX atherosclerosis; sickle cell anaemia; thalassaemia;

XX blood clotting disorder; diabetes; ss.

OS Adeno associated virus.

XX US6156303-A.

XX 05-DEC-2000.

XX 11-JUN-1997; 97US-0873168.

XX 11-JUN-1997; 97US-0873168.

PA (UNIW ) UNIV WASHINGTON.

XX Russell DW, Rutledge EA;

XX

|   |   |           |               |  |      |  |  |  |  |  |
|---|---|-----------|---------------|--|------|--|--|--|--|--|
| DR  | WPI; 2001-060164/07.  |           |               |  |      |  |  |  |  |  |
| XX  |   |           |               |  |      |  |  |  |  |  |
| PT  | Adeno-associated virus serotype 6 and viral vector derived from it for    |           |               |  |      |  |  |  |  |  |
| PT  | gene therapy of cystic fibrosis, cancer, acquired immunodeficiency        |           |               |  |      |  |  |  |  |  |
| PT  | syndrome, sickle cell anemia, thalassemia and diabetes -                  |           |               |  |      |  |  |  |  |  |
| XX  |   |           |               |  |      |  |  |  |  |  |
| PS  | Example 2; Fig 1; 50pp; English.  |           |               |  |      |  |  |  |  |  |
| XX  |   |           |               |  |      |  |  |  |  |  |
| CC  | The present invention relates to adeno-associated virus serotypes. The    |           |               |  |      |  |  |  |  |  |
| CC  | present sequence is the DNA sequence of one such serotype (AAV3B). AAV3B  |           |               |  |      |  |  |  |  |  |
| CC  | can be used to construct AAV viral vectors for use in gene therapy for a  |           |               |  |      |  |  |  |  |  |
| CC  | range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,       |           |               |  |      |  |  |  |  |  |
| CC  | sickle cell anaemia, thalassaemia, blood clotting disorders and diabetes. |           |               |  |      |  |  |  |  |  |
| CC  | The AAV viral vectors have increased transduction efficiency of a         |           |               |  |      |  |  |  |  |  |
| CC  | particular host cell as the AAV virion containing the AAV vector genome   |           |               |  |      |  |  |  |  |  |
| CC  | can be modified to express a capsid protein of an AAV serotype that       |           |               |  |      |  |  |  |  |  |
| CC  | transduces the selected host cell.  |           |               |  |      |  |  |  |  |  |
| XX  |   |           |               |  |      |  |  |  |  |  |
| SQ  | Sequence 4722 BP; 1204 A; 1297 C; 1243 G; 978 T; 0 other;                 |           |               |  |      |  |  |  |  |  |
| Alignment Scores:                             |   |           |               |  |      |  |  |  |  |  |
| Pred. No.:                                    |   | 7.76e-248 | Length:       |  | 4722 |  |  |  |  |  |
| Score:  |   | 3511.00   | Matches:      |  | 640  |  |  |  |  |  |
| Percent Similarity:                           |   | 91.99%    | Conservative: |  | 38   |  |  |  |  |  |
| Best Local Similarity:                        |   | 86.84%    | Mismatches:   |  | 57   |  |  |  |  |  |
| Query Match:                                  |   | 88.02%    | Indels:       |  | 2    |  |  |  |  |  |
| DB:   |   | 22        | Gaps:         |  | 2    |  |  |  |  |  |
| US-09-807-802A-13 (1-736) x AAF23748 (1-4722) |   |           |               |  |      |  |  |  |  |  |
| QY  | 1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20         |           |               |  |      |  |  |  |  |  |
| Db  | 2208 ATGGCTGCTGACGGTTATCTCCAGATTGGCTCGAGGACAACCTTCTCTGAAGGCATTTCGT 2267   |           |               |  |      |  |  |  |  |  |
| QY  | 21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp 40        |           |               |  |      |  |  |  |  |  |
| Db  | 2268 GAGTGGTGGGCTCTGAACCTGGAGTCCCTCAACCCAAAGCGAACCACACACAGGAC 2327        |           |               |  |      |  |  |  |  |  |
| QY  | 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp 60        |           |               |  |      |  |  |  |  |  |
| Db  | 2328 AACCGTCGGGGTCTTGTCCTCCGGTTACAAATACCTCGACCCCGGTAAACGGACTCGAC 2387     |           |               |  |      |  |  |  |  |  |
| QY  | 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80           |           |               |  |      |  |  |  |  |  |
| Db  | 2388 AAAGGAGAGCCGGTCAACGAGCGGACGCGGAGCCCTCGAACACGACAAAGCTTACGAC 2447      |           |               |  |      |  |  |  |  |  |
| QY  | 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100       |           |               |  |      |  |  |  |  |  |
| Db  | 2448 CAGCAGCTCAAGCCGGTGACAACCCGTACCTCAAGTACAAACCAACGCGGACGCGAGTTT 2507    |           |               |  |      |  |  |  |  |  |
| QY  | 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120      |           |               |  |      |  |  |  |  |  |
| Db  | 2508 CAGGAGCGTCTTCAAGAAGATACGTCTTTTGGGGCAACCTTGGCAGAGCAGTCTTCCAG 2567     |           |               |  |      |  |  |  |  |  |
| QY  | 121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro 140         |           |               |  |      |  |  |  |  |  |
| Db  | 2568 GCCAAAAGAGGATCCTTGAGCCTCTTGGTCTGGTTGAGGAAGCAGCTAAAACGGCTCCT 2627     |           |               |  |      |  |  |  |  |  |
| QY  | 141 GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly 160      |           |               |  |      |  |  |  |  |  |
| Db  | 2628 GGAAGAAGAGGCCCTGTAGATCAGTCTCCTCAGGAACCGGACTCATCATCTGGTGTGGC 2687     |           |               |  |      |  |  |  |  |  |
| QY  | 161 LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180      |           |               |  |      |  |  |  |  |  |
| Db  | 2688 AAATCGGGCAACAGCCCTGCCAGAAAAGACTAAATTTCCGTTCAGACTCGGCTCAGAG 2747      |           |               |  |      |  |  |  |  |  |
| QY  | 181 SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro 200      |           |               |  |      |  |  |  |  |  |
| Db  | 2748 TCAGTCCCAGACCCTCAACCTCTCGGAGAAACCACGACGACGCCCCCAAGTTTGGGATCT 2807    |           |               |  |      |  |  |  |  |  |
| QY  | 201 ThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly 220      |           |               |  |      |  |  |  |  |  |
| Db  | 2808 AATACAATGGCTTCAGCGGTCGCGCACCAATGGCAGACAAATAACGAGGTCGCCGATGGA 2867    |           |               |  |      |  |  |  |  |  |

|    |      |   |
|----|------|---|
| QY | 221  | ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValile 240    |
| Db | 2868 | GTGGTAAATTCCTCAGGAAATTGGCATTCGATTCCCAATGGCTGGCGCAGAGTCATC 2927      |
| QY | 241  | ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnile 260    |
| Db | 2928 | ACCACCAGCACCAGAACCTGGGCCCTGCCACTTACAACAACCATCTCTACAAGCAATC 2987     |
| QY | 261  | SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTip 280    |
| Db | 2988 | TCCAGCCCAATCA--GGAGCTTCAAACGACAACCACTACTTTGGCTACAGCACCCCTTGG 3044   |
| QY | 281  | GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu 300    |
| Db | 3045 | GGGTATTTTGACTTTAACAGATTCCACTGCCACTTCTCACCACGTGACTGGCAGCGACTC 3104   |
| QY | 301  | IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln 320    |
| Db | 3105 | ATTAACAACAACCTGGGGATTCCGGCCCCAAGAAACTCAGCTTCAAGCTCTTCAACATCCAA 3164 |
| QY | 321  | ValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThr 340    |
| Db | 3165 | GTAAAGAGGTCAAGACGATGGCAGCAGCACTATTGGCAATAAACCTTACCAGCAGC 3224       |
| QY | 341  | ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln 360    |
| Db | 3225 | GTTCAGTGTTTACGGACTCGGAGTATCAGCTCCCGTACGTCTCGGGTCGGCGCACCA 3284      |
| QY | 361  | GlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr 380       |
| Db | 3285 | GGCTGTCTCCGCCGTTTCCAGCGGACGTCTTCATGGTCCCTCAGTATGGATACCTCACC 3344    |
| QY | 381  | LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 400    |
| Db | 3345 | CTGAACAACCGAAGTCAAGCGGTGGACGCTCATCTTTTACTGCTCGCTGGAGTACTTCCCT 3404  |
| QY | 401  | SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro 420    |
| Db | 3405 | TCGCAGATGCTAAGGACTGGAAATAACTTCCAATTCACTATACCTTCGAGGATGTACCT 3464    |
| QY | 421  | PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 440    |
| Db | 3465 | TTTTCACAGCAGCTACGCTCACAGCCAGAGTTTGGATCGCTTGATGAATCCTCTTATTGAT 3524  |
| QY | 441  | GlnTyrLeuTyrTyrLeuAsnArgThrGln--AsnGlnSerGlySerAlaGlnAsnLys 459     |
| Db | 3525 | CAGTATCTGTACTACTCTGAACAGAACGCAAGGAACAACCTCTCGGAACAACCAACCAATCA 3584 |
| QY | 460  | AspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTipLeu 479    |
| Db | 3585 | CGGCTGCTTTTAGCCAGGCTGGGCCCTCAGTCTATGTCTTTGCAGGCCAGAAATTGGCTA 3644   |
| QY | 480  | ProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnSer 499       |
| Db | 3645 | CCTGGGCCCTGCTACCGGCAACAGAGACTTTCAAAGACTGCTAACGACAACAACACAGT 3704    |
| QY | 500  | AsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsn 519    |
| Db | 3705 | AACCTTCTTGACAGCGGCCAGCAATAATATCATCTCAATGGCCGCGACTCGTGGTGAAT 3764    |
| QY | 520  | ProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGly 539    |
| Db | 3765 | CCAGGACCAGCTATGGCCAGTCAACAAGGACGATGAAGAAAATTTTCCCTATGCACGGC 3824    |
| QY | 540  | ValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMet 559    |
| Db | 3825 | AATCTAATATTGGCAAGAAGGGACACCGGCAAGTAACGCAGAAATTAGATAATGTAATG 3884    |
| QY | 560  | IleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThr 579       |
| Db | 3885 | ATTACGGATGAAGAAGAGATTTCGTACCACCAATCCTGTGGCAACAGACGATATGGAAC 3944    |





|   |        |   |      |
|---|--------|---|------|
| Query Match:                                  | 86.21% | Indels:   | 2    |
| DB:   | 21     | Gaps:   | 2    |
| US-09-807-802A-13 (1-736) x AAD00834 (1-8178) |        |   |      |
| QY  | 1      | MetAlaAlaaspGlyTyrLeuProAspTrpLeuGluAaspAsnLeuSerGluGlyIleArg | 20   |
| Db  | 2133   | ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACACTCTCTCTGAAGGAATAAGA  | 2192 |
| QY  | 21     | GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp  | 40   |
| Db  | 2193   | CAGTGGTGGAAAGCTCAAACTGGCCCAACCAACCAAGCCCGCAGAGCGGCATAAGGAC    | 2252 |
| QY  | 41     | AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp  | 60   |
| Db  | 2253   | GACAGCAGGGGTCTTGTGCTTCTCTGGGTACAAGTACCTCGGACCTTCAACGGACTCGAC  | 2312 |
| QY  | 61     | LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp     | 80   |
| Db  | 2313   | AAGGAGAGCCGCTCAACGAGGACAGCGCGCGGCCCTCGAGCACGACAAAAGCCTACGAC   | 2372 |
| QY  | 81     | GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe  | 100  |
| Db  | 2373   | CGGCAGCTCGACAGCGGAGACAACCCGTACTCAAGTACAACCAACCGCGACGCGGAGTTT  | 2432 |
| QY  | 101    | GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln  | 120  |
| Db  | 2433   | CAGGAGCGCCTTAAAGAAGATACGTCCTTTGGGGCAACCTCGGAGCAGCAGCTTCCAG    | 2492 |
| QY  | 121    | AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro     | 140  |
| Db  | 2493   | GCGAAAAAGAGGGTCTTTGAACCTCTGGGCTGGTTGAGGAACCTCTCTCGGAAACCGGA   | 2552 |
| QY  | 141    | GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerGlyIleGly     | 160  |
| Db  | 2553   | GGAAAAAAGAGCGCGGTAGAGCACTCTCTGTGGAGCCAGACTCTCTCTCGGAAACCGGA   | 2612 |
| QY  | 161    | LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu  | 180  |
| Db  | 2613   | AAGCGGGCCAGCAGCCTGCAAGAAAAAGATTGAATTTTGGTCAGACTGGAGACGCAGAC   | 2672 |
| QY  | 181    | SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro  | 200  |
| Db  | 2673   | TCAGTACCTGACCCCGCAGCCTCTCGGACAGCCACCCAGCAGCCCCCTCTGGTCTGGGA   | 2732 |
| QY  | 201    | ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly     | 220  |
| Db  | 2733   | AATACGATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGCGCGGACCGA   | 2792 |
| QY  | 221    | ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle  | 240  |
| Db  | 2793   | GTGGGTAATTCTTCGGAAATTGGCATTTGCATTTCCATGGATGGGCGACAGATCATC     | 2852 |
| QY  | 241    | ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle  | 260  |
| Db  | 2853   | ACCACCAGCACCCGAACCTGGGCCCTGCCCACTTACAAACCAACCACTCTACAAACAA    | 2912 |
| QY  | 261    | SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp  | 280  |
| Db  | 2913   | TCCAGCCAATCA---GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTG   | 2969 |
| QY  | 281    | GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu  | 300  |
| Db  | 2970   | GGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCAACACGCTGAGTGGCAAGACTC  | 3029 |
| QY  | 301    | IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln  | 320  |
| Db  | 3030   | ATCAACAACAACCTGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAACA      | 3089 |
| QY  | 321    | ValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThr  | 340  |
| Db  | 3090   | GTCAAAGAGGTCACGCAGAAATGACGGTACGAGCAGGATTGCCAATAAACCTTACCAG    | 3149 |
| QY  | 341    | ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln  | 360  |
| Db  | 3150   | GTTCAGGTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTGCTCGGGTCGGCGACCAA   | 3209 |
| QY  | 361    | GlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr  | 380  |
| Db  | 3210   | GGCTGTCTCCCGCCGTTTCCAGCGGACGTCTTTCATGGTCCCTCAGTATGGATACCTCAC  | 3269 |
| QY  | 381    | LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro  | 400  |
| Db  | 3270   | CTGAACAACGGAAGTCAAGCGGTGGGACGCTCATCTTTTACTGCTGGAGTACTTCCCT    | 3329 |
| QY  | 401    | SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro  | 420  |
| Db  | 3330   | TCGCAGATGCTAAGGACTGGAAATAAATTCCAATTTCAGCTATACCTTCGAGGATGTACT  | 3389 |
| QY  | 421    | PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp  | 440  |
| Db  | 3390   | TTTCACAGCAGCTACGCTCACAGCCAGAGTTTGGATCGCTTGATGAATCCTCTTATTGAT  | 3449 |
| QY  | 441    | GlnTyrLeuTyrTyrLeuAsnArgThrGln---AsnGlnSerGlySerAlaGlnAsnLys  | 459  |
| Db  | 3450   | CAGTATCTGTACTACTCCTGAACAGAACGCAAGGAACAACCTCTGGAACAACAACCAATCA | 3509 |
| QY  | 460    | AspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeu  | 479  |
| Db  | 3510   | CGGCTGCTTTTAGCCAGGCTGGGCCTCAGTCTATGTCTTTCGAGGCCAGAAATTGGCTA   | 3569 |
| QY  | 480    | ProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSer  | 499  |
| Db  | 3570   | CCTGGGCCCTGTACTACCGCAACAGAGACTTTTCAAAGACTGCTAAACGACCAACAACAGT | 3629 |
| QY  | 500    | AsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleAsn     | 519  |
| Db  | 3630   | AACTTTCCTTGGACAGCGGCAGCAAAATATCATCTCAATGGCCGCGACTCGTGGTGAAT   | 3689 |
| QY  | 520    | ProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGly  | 539  |
| Db  | 3690   | CCAGGACCAGCTATGGCCAGTCAACAAGGACGATGAAGAAATAATTTTCCCTATGCACG   | 3749 |
| QY  | 540    | ValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMet  | 559  |
| Db  | 3750   | AATCTAATAATTTGGCAAGAAGGGACAACCGCAAGTAACGAGAAATTAGATAATGTATG   | 3809 |
| QY  | 560    | IleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThr  | 579  |
| Db  | 3810   | ATTACGGATGAAGAAGAGATCGTACCACCAATCCTGTGGCAACAGAGCAGTATGGAACT   | 3869 |
| QY  | 580    | ValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMet  | 599  |
| Db  | 3870   | GTGGCAAAATAACTTGCAGAGCTCAAATACAGCTCCCAAGACTGGAATGTCAATCATCAG  | 3929 |
| QY  | 600    | GlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrp  | 619  |
| Db  | 3930   | GGGGCTTACCTGGCATGGTGGCAAGATCGTGACGTGTACCTTCAAGGACCTATCTGG     | 3989 |
| QY  | 620    | AlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGly  | 639  |
| Db  | 3990   | GCAAGATTCTCTCACACGGATGGACACTTTCATCTCTCTCTCTGTATGGAGGCTTTTGA   | 4049 |
| QY  | 640    | LeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProPro     | 659  |
| Db  | 4050   | CTGAACAATCCGCCCTCCTCAAATCATGATCAAAAATACTCCGGTACCTGCGAATCCTTCG | 4109 |
| QY  | 660    | AlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnVal  | 679  |
| Db  | 4110   | ACCACCTTCAGTCCGGCAAGTTTGTCTTCTCATCACAGTACTCCACGGGACAGGTC      | 4169 |
| QY  | 680    | SerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluVal  | 699  |
| Db  | 4170   | AGCCTGGAGATCGAGTGGAGCTGCAGNAGGAAACAGCAAAACGCTGGAATCCCGAAAT    | 4229 |
| QY  | 700    | GlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGly  | 719  |

Db 4230 CAGTACACTTCCAACACTACAAGCTCTGTTAATCGTGGACTTACCGTGGATACTAATGGC 4289  
QY 720 LeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736  
Db 4290 GTGTATTTCAGAGCCTCGCCCCATTGGCACCAGATACCTGACTCGTAATCTG 4340

RESULT 6  
AAH26327  
ID AAH26327 standard; DNA; 4072 BP.  
XX  
AC AAH26327;  
DT 02-OCT-2001 (first entry)  
XX  
DE Adenovirus helper Ad cap2.  
XX  
KW AAV; vector; adenovirus; helper virus; Ad cap2; loxP site;  
KW gene therapy; ds.  
XX  
OS Chimeric - Mastadenovirus.  
OS Chimeric - Adeno associated virus.  
OS Chimeric - Human cytomegalovirus.  
XX  
PN WO200155361-A2.

XX 02-AUG-2001.  
XX  
XX 26-JAN-2001; 2001WO-US02709.  
XX  
XX 26-JAN-2000; 2000US-0178536.  
XX (CHIR ) CHIRON CORP.  
XX  
PI Hardy SF;  
XX  
XX WPI; 2001-483239/52.  
XX  
PT Producing recombinant adeno-associated virus (rAAV) vector, by stably  
PT infecting eukaryotic host cell with rAAV vectors, helper adenovirus or  
PT herpes virus, recombinant adenovirus or herpes vectors -

PS Example 2; Page 58-59; 63pp; English.  
XX  
CC The present sequence is that of adenovirus helper Ad cap2, made  
CC by Cre/lox recombination and expressing adeno-associated virus  
CC (AAV) VP1, 2,3 from the cytomegalovirus immediate early promoter  
CC of pAdlox, and containing a loxP site. Ad cap2 was used to  
CC demonstrate recombinant AAV (rAAV) production from virally  
CC transduced cells. The invention provides methods and compositions  
CC for producing rAAV vector particles by: (a) introducing into a host  
CC cell (i) AAV packaging plasmid pfloxAAV (see AAH26326), (ii) a  
CC recombinant viral vector encoding plasmid, and (iii) a plasmid  
CC encoding herpesvirus, cytomegalovirus or adenoviral functions, or  
CC a herpesvirus, cytomegalovirus or adenovirus itself, to produce  
CC flox AAV particles and rAAV particles; and (b) introducing into a  
CC second host cell (i) the rAAV particles or (a), (ii) a vector  
CC that directs expression of Cre, and (ii) a vector which directs  
CC expression of herpes virus, cytomegalovirus or adenovirus helper  
CC functions, such that rAAV vector particles are produced. The  
CC vectors are useful for in vivo or in vitro gene therapy and also  
CC for in vitro recombinant protein production.

XX  
SQ Sequence 4072 BP; 1073 A; 1025 C; 992 G; 982 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 5.95e-240 Length: 4072  
Score: 3402.50 Matches: 613  
Percent Similarity: 90.22% Conservative: 51  
Best Local Similarity: 83.29% Mismatches: 71  
Query Match: 85.30% Indels: 1  
DB: 22 Gaps: 1

US-09-807-802A-13 (1-736) x AAH26327 (1-4072)  
QY 1 MetAlaAlaAspGlyTyrLeuProAspTyrLeuGluAspAsnLeuSerGluGlyIleArg 20  
Db 1484 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACACTCTCTCTGAAGGAATAAGA 1543  
QY 21 GluTyrTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp 40  
Db 1544 CAGTGGTGAAGCTCAAAACCTGGCCACCACCAACCAAGCCCGCAGAGCGGATAGGAC 1603  
QY 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp 60  
Db 1604 GACAGCAGGGGTCTTGTGTTCTCTGGGTACAAGTACCTCGGACCTTCAACGGACTCGAC 1663  
QY 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80  
Db 1664 AAGGAGAGCGGTCAACGAGGCAGACGCGCGGCTCGAGCAGCAAGCCCTACGAC 1723  
QY 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100  
Db 1724 CGGCAGCTCGACAGCGGAGACAACCCGTACCTCAAGTACAACCCAGCCGCGAGTTT 1783  
QY 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120  
Db 1784 CAGGAGCGCCTTAAAGAAGATACGTCTTTTGGGGCAACCTCGGACGAGCAGTCTTCCAG 1843  
QY 121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro 140  
Db 1844 GCGAAAAAGAGGGTTCTTGAACCTCTGGGCTGTTGAGGAACCTGTTAAGACGGCTCCG 1903  
QY 141 GlyLysLysArgProValGlnSerProGlnGluProAspSerSerSerGlyIleGly 160  
Db 1904 GGAATAAAGAGGCGGTAGAGCACTCTCTCTGTGGAGCAGACTCTCTCTCGGGAACCGGA 1963  
QY 161 LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180  
Db 1964 AAGCGGGCCAGCAGCGCTGCAAGAAAAGATTGAATTTGGTCAGACTGGAGACGCAGAC 2023  
QY 181 SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro 200  
Db 2024 TCAGTACCTGACCCCCAGCCTCTCGGACAGCCACCAGAGCCCCCTCTGGTCTGGGAAC 2083  
QY 201 ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnGluGlyAlaAspGly 220  
Db 2084 AATACGATGGCTACAGGCAGTGGCGCACCAATGGAGACAATAACGAGGGCGCCGACGGA 2143  
QY 221 ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle 240  
Db 2144 GTGGTAAATTCCTCGGNAATTGGCATTCGATTCACATGGATGGGCGACAGAGTCATC 2203  
QY 241 ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle 260  
Db 2204 ACCACGACACCGAACCTGGGCTGTCACCACTACAAACACCACTCTACAAACAAAT 2263  
QY 261 SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp 280  
Db 2264 TCCAGCCAATCA---GGAGCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTG 2320  
QY 281 GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu 300  
Db 2321 GGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCCACCGTGTGGCAAAGACTC 2380  
QY 301 IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln 320  
Db 2381 ATCAACAACAACCTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAATCA 2440  
QY 321 ValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThr 340  
Db 2441 GTCAAGAGGTCAACGAGATGACGGTACGACGACGATTGCCAATAAATTAACCTTACCAG 2500  
QY 341 ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln 360  
Db 2501 GTTCAGGTGTTTACTGACTCGAGTACCAGCTCCCGTACGTCTCGGCTCGGCGCATCAA 2560

QY 361 GlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr 380  
Db 2561 GGATGCTCCCGCGTTCGCCAGCAGCGTCTTCATGGTGCCACAGTATGGATACCTCACC 2620

QY 381 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 400  
Db 2621 CTGAACAACGGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCTGGAGTACTTTCCT 2680

QY 401 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro 420  
Db 2681 TCTCAGATGCTGCGTACCGGAACAACACTTTACCTTCAGCTACACTTTTGAGGACGTTCT 2740

QY 421 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 440  
Db 2741 TTCCACAGCAGCTACGCTCACAGCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGAC 2800

QY 441 GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAsp 460  
Db 2801 CAGTACCTGTATTACTTGAGCAGACAACAACACTCCAAGTGGAAACCAACCACGCGTCAAGG 2860

QY 461 LeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuPro 480  
Db 2861 CTTTCAGTTTCTCAGGCCGCGAGCGAGTGCATTCGGGACCACTCTAGGAACCTGCTTCCT 2920

QY 481 GlyProCysTyrArgGlnGlnArgValSerLysThrLysThrLysThrAspAsnAsnSerAsn 500  
Db 2921 GGACCCCTGTACCGCCAGCAGCGAGTATCAAGACATCTCGGATACAACAACAGTGA 2980

QY 501 PheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnPro 520  
Db 2981 TACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCG 3040

QY 521 GlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyVal 540  
Db 3041 GGCCCGGCATGGCAAGCCACAGGAGCATGAAGAAAGTTTTTCTCTCAGAGCGGGGT 3100

QY 541 MetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIle 560  
Db 3101 CTCATCTTTGGGAAGCAAGGCTCAGAGAAAACAAATGTGGACATGTGAAGGTCATGATT 3160

QY 561 ThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrVal 580  
Db 3161 ACAGACGAAGAGGAAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTATGTTCTGTA 3220

QY 581 AlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGly 600  
Db 3221 TCTACCAACCTCCAGAGAGGCAACAGCAAGCAGCTACCGCAGATGTCAACACACAAGGC 3280

QY 601 AlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAla 620  
Db 3281 GTTCTTCCAGGATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGCA 3340

QY 621 LysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeu 640  
Db 3341 AAGATTCCACACGCGACGGACGACATTTTCACCCCTCTCCCTCATGGGTGGATTTCGGACT 3400

QY 641 LysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAla 660  
Db 3401 AAACACCCCTCTCCACAGATTCTCATCAAGAACACCCCGGTACCTCGGAATCCTTCGACC 3460

QY 661 GluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSer 680  
Db 3461 ACCTTCAGTGGCGAAAAGTTTGCTTCCTTCATCACACAGTACTCCACGGGACAGGTCCAGC 3520

QY 681 ValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGln 700  
Db 3521 GTGGAGATCGAGTGGAGCTGCAGAGGAAAACAGCAACGCTGGAATCCCGAATTCAG 3580

QY 701 TyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeu 720  
Db 3581 TACACTTCCAACCTACAACAGTCTGTTAATGTGGACTTTACTGTGGACACTAATGGCGTG 3640

QY 721 TyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736  
Db 3641 TATTGAGAGCCTCGCCCATTTGGCCACCAGATACCTGACTCGTAATCTG 3688

RESULT 7  
AAI66974  
ID AAI66974 standard; DNA; 4679 BP.  
XX AAI66974;  
AC AAI66974;  
XX 11-FEB-2002 (first entry)  
XX Adeno-associated virus 2 complete genomic sequence.  
XX Recombinant viral vector; RWV; capsid; parvovirus; transgene; cytotstatic;  
KW inverted terminal repeat; nootropic; neuroprotective; antianemic; ITR;  
KW antidiabetic; antitumour; gene therapy; adeno-associated virus; AAV; ds.  
XX Adeno-associated virus 2.  
XX Key Location/Qualifiers  
CDS 321..2252  
FT /\*tag= a  
FT /product= "Rep 68 protein"  
FT /note= "contains introns; for coding sequence  
join (321..1906, 2228..2252)"  
FT CDS 321..2186  
FT /\*tag= b  
FT /product= "Rep 78 protein"  
FT 993..2252  
FT /\*tag= c  
FT /product= "Rep 40 protein"  
FT /note= "contains introns; for coding sequence  
join (993..1906, 2228..2252)"  
FT CDS 993..2186  
FT /\*tag= d  
FT /product= "Rep 52 protein"  
FT 2203..4410  
FT /\*tag= e  
FT /product= "major coat protein VP1"  
FT 2614..4410  
FT /\*tag= f  
FT /transl\_except= "(pos:2614..2616, aa: Met)"  
FT /product= "major coat protein VP2"  
FT 2809..4410  
FT /\*tag= g  
FT /product= "major coat protein VP3"  
XX WO200168888-A2.  
PN 20-SEP-2001.  
XX 13-MAR-2001; 2001WO-US07927.  
XX 14-MAR-2000; 2000US-189110P.  
XX (NEUR-) NEUROLOGIX INC.  
XX Xiao W, During MJ;  
XX WPI; 2001-596912/67.  
DR P-PSDB; AAG65788, AAG65789, AAG65790, AAG65791, AAG65792, AAG65793,  
DR AAG65794.  
XX Recombinant viral vector useful in improving gene therapy in a subject,  
PT and for increasing efficiency of entry into a cell, comprises a  
PT chimeric capsid having one non-native amino acid sequence and a desired  
PT transgene  
XX Disclosure; Page 46-47; 53pp; English.  
PS The invention provides a recombinant viral vector (RWV) comprising a  
XX chimeric capsid (I) having at least one non-native amino acid sequence,  
CC



CC derived from a capsid protein domain of parvovirus (II), a virus (III),  
CC or their combination, and a transgene flanked 5' and 3' by inverted  
CC terminal repeat (ITR) sequences, derived from (II), (III), or their  
CC combination. The RVV is useful for improving gene therapy in a subject  
CC with a disorder, and for increasing the efficiency of entry into a cell,  
CC which involves producing (I) encapsulating a viral vector, and contacting  
CC a cell with RVV having (I) such that (I) binds to an attachment site on  
CC the cell surface and permits the vector to enter the cell efficiently.  
CC A pharmaceutical composition comprising RVV with (I) containing a  
CC transgene sequence associated with a disease or a disorder such that  
CC expression of the transgene would result in amelioration of the disease  
CC or disorder such as inherited neurological and metabolic diseases e.g.  
CC lysosomal storage disease, Lesch-Nyhan syndrome, amyloid polynuropathy,  
CC Alzheimer amyloid, Duchenne's muscular dystrophy, diseases of the blood  
CC e.g. sickle-cell anemia, clotting disorders and thalassemias, cystic  
CC fibrosis, diabetes, diseases associated with hormone deficiencies,  
CC retinoblastoma and various types of neoplastic cells which include  
CC tumours especially central nervous system tumours, neoplasms, carcinomas,  
CC sarcomas, leukemias and lymphoma. The present sequence represents the  
CC complete genomic sequence of adeno-associated virus 2 which is used in  
CC the construction of a chimeric vector.

XX  
SQ Sequence 4679 BP; 1198 A; 1262 C; 1255 G; 964 T; 0 other;

Alignment Scores:

|                        |          |               |      |
|------------------------|----------|---------------|------|
| Pred. No.:             | 7.1e-240 | Length:       | 4679 |
| Score:                 | 3402.50  | Matches:      | 613  |
| Percent Similarity:    | 90.22%   | Conservative: | 51   |
| Best Local Similarity: | 83.29%   | Mismatches:   | 71   |
| Query Match:           | 85.30%   | Indels:       | 1    |
| DB:                    | 22       | Gaps:         | 1    |

US-09-807-802A-13 (1-736) x AAI66974 (1-4679)

|    |      |  |      |
|----|------|--|------|
| QY | 1    | MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg | 20   |
| Db | 2203 | ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACACTCTCTCTGAAGGAATAAGA | 2262 |
| QY | 21   | GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnLysGlnAsp    | 40   |
| Db | 2263 | CAGTGGTGAAGCTCAACCTGGCCCAACCACCACCAAGCCGCGAGCGGCATAAGGAC     | 2322 |
| QY | 41   | AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyPropheAsnGlyLeuAsp | 60   |
| Db | 2323 | GACAGCAGGGTCTTGTTCTCTGGGTACAAGTACCTCGGACCTTCAACGGACTCGAC     | 2382 |
| QY | 61   | LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp    | 80   |
| Db | 2383 | AAGGGAGAGCGGTCAACGAGGACGACGCGCGGCCCTCGAGCAGCACAAGCCTACGAC    | 2442 |
| QY | 81   | GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe | 100  |
| Db | 2443 | CGGCAGCTCGACAGCGGAGACAAACCCGTACCTCAAGTACAACCAACCGCGGAGTTT    | 2502 |
| QY | 101  | GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln | 120  |
| Db | 2503 | CAGGAGCGCCTTAAAGAAGATACGTCCTTGGGGGCAACCTCGGACGAGCAGTCTTCCAG  | 2562 |
| QY | 121  | AlaLysLysArgValLeuGluProLeuGlyLeuValGluGluGlyAlaLysThrAlaPro | 140  |
| Db | 2563 | GCGAAAAGAGGGTCTTGAACCTCTGGGCTGGTTGAGGAACCTGTTAAGACGGCTCCG    | 2622 |
| QY | 141  | GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerGlyIleGly    | 160  |
| Db | 2623 | GGAAAAAGAGCGCGGTAGAGCACTCTCTCTGGGAGCCAGACTCTCTCTCGGGAAACCGGA | 2682 |
| QY | 161  | LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu | 180  |
| Db | 2683 | AAGCGCGGCAGCAGCCTGCAAGAAAAAGATTGAATTTTGGTCAGACTGGAGACGCAGAC  | 2742 |
| QY | 181  | SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro | 200  |
| Db | 2743 | TCAGTACCTGACCCCGCCTCTCTCGGACAGCCACCAGCAGCCCCCTCTGGTCTGGGAAC  | 2802 |

|    |      |  |      |
|----|------|--|------|
| QY | 201  | ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnGluGlyAlaAspGly         | 220  |
| Db | 2803 | AATACGATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGGCGCGACGGA    | 2862 |
| QY | 221  | ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValile   | 240  |
| Db | 2863 | GTGGGTAAATTCCTCGGAAATTGGCATTGGCATTCACATGGATGGGCGACAGAGTCATC    | 2922 |
| QY | 241  | ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle   | 260  |
| Db | 2923 | ACCACGAGCACCCGAACTGGGCCCTGCCACCTACAAACAACCAACCTCTACAAACAAT     | 2982 |
| QY | 261  | SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp   | 280  |
| Db | 2983 | TCCAGCCAATCA---GGAGCCTCGAACGACAACTCACTTTTGGCTACAGCACCCCTGG     | 3039 |
| QY | 281  | GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu   | 300  |
| Db | 3040 | GGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCCACCAGTGTGCTGCAAGACTC    | 3099 |
| QY | 301  | IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln   | 320  |
| Db | 3100 | ATCAACAACAACCTGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAACAATCAA  | 3159 |
| QY | 321  | ValLysGluValThrThrAsnAspGlyValThrIleAlaAsnAsnLeuThrSerThr      | 340  |
| Db | 3160 | GTCAAAGAGGTACGCAGAATGACGGTACGACGAGATTGCCAATAACCTTACCAGCAGC     | 3219 |
| QY | 341  | ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln   | 360  |
| Db | 3220 | GTTCAGGTGTTTACTGACTCGGAGTACCAGTCCCGTACGTCTCGGCTCGGCGCATCAA     | 3279 |
| QY | 361  | GlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr   | 380  |
| Db | 3280 | GGATGCCTCCCGCGTTCCAGCAGACGCTCTTCATGGTGCCACAGTATGGATACCTCACC    | 3339 |
| QY | 381  | LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro   | 400  |
| Db | 3340 | CTGAACAACGCGGAGTCAGGCAGTAGGACGCTCTTCACTTTTACTGCTGGAGTACTTCT    | 3399 |
| QY | 401  | SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro   | 420  |
| Db | 3400 | TCTCAGATGCTGCGTACCGGAACAACCTTTACCTTCAGCTACACTTTTGAGGACGTTCT    | 3459 |
| QY | 421  | PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp   | 440  |
| Db | 3460 | TTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACGCTCTCATGAATCCTCTCATCGAC   | 3519 |
| QY | 441  | GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAsp   | 460  |
| Db | 3520 | CAGTACCTGTATTACTTGAGCAGAAACAACACTCCAAGTGGAAACCAACCAGCAGTCAAGG  | 3579 |
| QY | 461  | LeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuPro   | 480  |
| Db | 3580 | CTTCAGTTTCTCAGCGCGGAGCGAGTGACATTCCGGGACCAGTCTAGGAACCTGGCTTCT   | 3639 |
| QY | 481  | GlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnSerAsn      | 500  |
| Db | 3640 | GGACCCCTGTTACCGCCAGCAGCGAGTATCAAAAGACATCTGCGGATAACAACAACAGTGAA | 3699 |
| QY | 501  | PheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnPro   | 520  |
| Db | 3700 | TACTCGTGAGCTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGTTGATCCG    | 3759 |
| QY | 521  | GlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyVal   | 540  |
| Db | 3760 | GGCCCGGCCCATGGCAAGCCCAAGGACGATGAAGAAAAGTTTTTCTCTCAGAGCGGGGTT   | 3819 |
| QY | 541  | MetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIle   | 560  |
| Db | 3820 | CTCATCTTTGGGAAGCAAGGCTCAGAGAAAACAATGTGGACATTGAAAAGGTCTATGATT   | 3879 |

QY 561 ThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrVal 580  
|||||  
Db 3880 ACAGACGAAGAGGAAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTATGGTCTGTA 3939  
581 AlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGly 600  
::: ||| |||  
Db 3940 TCTACCAACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAACACACAAGGC 3999  
601 AlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAla 620  
|||||  
Db 4000 GTTCTCCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTCAGGGGCCCATCTGGGCA 4059  
621 LysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeu 640  
|||||  
Db 4060 AAGATTCCACACACGCGACGACATTTTCACCCCTCTCCCTCATGGTGGATTGGGACTT 4119  
641 LysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAla 660  
|||:::|||||  
Db 4120 AAACACCCCTCTCCACAGATTCTCATCAAGAAACACCCCGGTACCTGCGAATCCTTCGACC 4179  
661 GluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSer 680  
|||||  
Db 4180 ACCTTCAGTGGGCAAGTTTGCTTCCTTCATCACACAGTACTCCACGGGACAGGTCAGC 4239  
681 ValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGln 700  
|||||  
Db 4240 GTGGAGATCGAGTGGGAGCTGCAGAAGGAAACAGCAACCGTGGAAATCCCGAAATTCAG 4299  
701 TyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeu 720  
4300 TACACTTCCAACATACAACAAGTCTGTAAATGTGGACTTTTACTGTGGACACTAATGGCGTG 4359  
721 TyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736  
4360 TATTCAGAGCCTGCCCCATTGGCACAGATACCTGACTCGTAATCTG 4407

RESULT 8

AAAF23750  
ID AAF23750 standard; DNA; 4679 BP.  
XX  
AC AAF23750;  
XX  
DT 28-MAR-2001 (first entry)  
XX  
DE AAV2 DNA sequence.  
XX  
KW AAV2; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;  
KW atherosclerosis; sickle cell anaemia; thalassaemia;  
KW blood clotting disorder; diabetes; ss.  
XX  
OS Adeno associated virus.  
XX  
PN US6156303-A.  
XX  
PD 05-DEC-2000.  
XX  
PF 11-JUN-1997; 97US-0873168.  
XX  
PR 11-JUN-1997; 97US-0873168.  
XX  
PA (UNIW ) UNIV WASHINGTON.  
XX  
PI Russell DW, Rutledge EA;  
XX  
DR WPI; 2001-060164/07.  
XX  
PT Adeno-associated virus serotype 6 and viral vector derived from it for  
PT gene therapy of cystic fibrosis, cancer, acquired immunodeficiency  
PT syndrome, sickle cell anemia, thalassemia and diabetes -  
XX  
PS Claim 7; Fig 1; 50pp; English.  
XX  
CC The present invention relates to adeno-associated virus serotypes. The

CC present sequence is the DNA sequence of one such serotype (AAV2). AAV2  
CC can be used to construct AAV viral vectors for use in gene therapy for a  
CC range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,  
CC sickle cell anaemia, thalassaemia, blood clotting disorders and diabetes.  
CC The AAV viral vectors have increased transduction efficiency of a  
CC particular host cell as the AAV virion containing the AAV vector genome  
CC can be modified to express a capsid protein of an AAV serotype that  
CC transduces the selected host cell.

XX  
SQ Sequence 4679 BP; 1197 A; 1263 C; 1255 G; 964 T; 0 other;

Alignment Scores:

Pred. No.: 7.1e-240 Length: 4679  
Score: 3402.50 Matches: 613  
Percent Similarity: 90.22% Conservative: 51  
Best Local Similarity: 83.29% Mismatches: 71  
Query Match: 85.30% Indels: 1  
DB: 22 Gaps: 1

US-09-807-802A-13 (1-736) x AAF23750 (1-4679)

QY 1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20  
Db 2203 ATGGCTGCCGATGGTTATCTTCAGATTGGTCTGAGGACACTCTCTCTGAAGGAATAAGA 2262  
QY 21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp 40  
Db 2263 CAGTGGTGAAGCTCAAACCTGGCCACCACCAACCAAGCCCGCAGACGCGGATAAGGAC 2322  
QY 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyPropheAsnGlyLeuAsp 60  
Db 2323 GACAGCAGGGGTCTTGTCTTCTGGGTACAAGTACCTCGGACCTTCAACGCGACTCGAC 2382  
QY 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80  
Db 2383 AAGGGAGAGCCGTCAACGAGGCAGACGCGCGGCCCTCGAGCAGCAAAAGCCCTACGAC 2442  
QY 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100  
Db 2443 CGGCAGCTCGACAGCGGAGACAACCCGTACCTCAAGTACAACCCAGCCGCGAGATT 2502  
QY 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120  
Db 2503 CAGGAGCGCCTTAAAGAAGATACGTCTTTTGGGGCAACCTCGGACGAGCAGTCTTCCAG 2562  
QY 121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro 140  
Db 2563 GCGAAAAAGAGGGTTCTTGAACCTCTGGGCCTGGTTGAGGAACCTGTTAAGACGGCTCCG 2622  
QY 141 GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly 160  
Db 2623 GGAAAAAAGAGGCCGCTAGAGCACTCTCCTGTGGAGCAGACTCCTCCTCGGGAACCGGA 2682  
QY 161 LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180  
Db 2683 AAGCGGGCCAGCAGCCTGCAAGAAAAGATTGAATTTGGTCAGACTGGAGACGCGAGAC 2742  
QY 181 SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro 200  
Db 2743 TCAGTACCTGACCCCCAGCCTCTCGGACAGCCACCAATGGCAGACAATAACGAGGGCGCCGCGGA 2802  
QY 201 ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly 220  
Db 2803 AATACGATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGGCGCCGCGGA 2862  
QY 221 ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle 240  
Db 2863 GTGGGTAATTCTCCGGAATTGGCATTTGCGATTCCCATGGATGGCGACAGAGTCATC 2922  
QY 241 ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle 260  
Db 2923 ACCACGAGCACCCGGAACCTGGGCCCTGCCCCACCTACACAAACCACTCTACAAACAATT 2982

QY 261 SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrrp 280  
Dbb TCCAGCCCAATCA--GGAGCCTCGAACGACAATCACTACTTTGGGTACAGCACCCCTTGG 3039  
QY 281 GlyTyrPheAspPheHisCysHisPheSerProArgAspTrrpGlnArgLeu 300  
Dbb 3040 GGGTATTTTGACTTCAACAGATTCCCACTGCCACTTTTCACCAAGTACTGGCAAGACTC 3099  
QY 301 IleAsnAsnAsnTrrpGlyPheArgProLysArgLeuAsnPhelysLeuPheAsnIleGln 320  
Dbb 3100 ATCAACAACAACACTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAACAATTCAA 3159  
QY 321 vallysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThr 340  
Dbb 3160 GTCAAAGAGGTACCGCAGAGATGACGGTACGACGACGATTGCCAATAAACCTTACCAGCAG 3219  
QY 341 valGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln 360  
Dbb 3220 GTTCAGGTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTCCTCGGCTCGGCGCATCAA 3279  
QY 361 GlyCysLeuProPropheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr 380  
Dbb 3280 GGATGCCTCCCGCGTTCCAGCAGACGCTCTTCATGGTGGCCACAGTATGGATACCTCACC 3339  
QY 381 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 400  
Dbb 3340 CTGAACAACGGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCTGGAGTACTTTTCCT 3399  
QY 401 SerGlnMetLeuArgThrGlyAsnAsnPhethePheSerTyrThrPheGluGluValPro 420  
Dbb 3400 TCTCAGATGCTGGGTACCGGAACAACATTACCTTCAGCTACACTTTTGAGGACGTTTCCT 3459  
QY 421 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 440  
Dbb 3460 TTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGAC 3519  
QY 441 GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAsp 460  
Dbb 3520 CAGTACCTGTATTACTTGAGCAGACAACAACACTCCAAGTGAACCAACCAAGCAGTCAAGG 3579  
QY 461 LeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrrpLeuPro 480  
Dbb 3580 CTTTCAGTCTTCTCAGCGCGGAGCGAGTGACATTTCGGGACCAGTCTAGGAACCTGGCTTCCT 3639  
QY 481 GlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsn 500  
Dbb 3640 GGACCCCTGTTACCGCCAGCAGCGAGTATCAAAAGACATCTCGGATATAACAACAAGTGAA 3699  
QY 501 PheThrTrrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnPro 520  
Dbb 3700 TACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCG 3759  
QY 521 GlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyVal 540  
Dbb 3760 GGCCCGGCCCATGGCAAGCCACAAGGACGATGAAGAAAGTTTTTTCTCAGAGCGGGGTT 3819  
QY 541 MetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIle 560  
Dbb 3820 CTTCATCTTTGGGAAGCAAGGCTCAGAGAAAACAATGTGGACATTGGAAGGTCATGATT 3879  
QY 561 ThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrVal 580  
Dbb 3880 ACAGACGAAGAGGAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTATGTTCTGTA 3939  
QY 581 AlaValAsnPhetheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGly 600  
Dbb 3940 TCTACCAACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAACAAGGC 3999  
QY 601 AlaLeuProGlyMetValTrrpGlnAspArgAspValTyrLeuGlnGlyProIleTrrpAla 620  
Dbb 4000 GTTCTTCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGCA 4059  
QY 621 LysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeu 640

Dbb 4060 AAGATTCCACACACGAGCGGACATTTTTCACCCCTCTCCCTCATGGGTGGATTCCGACTT 4119  
QY 641 LysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAla 660  
Dbb 4120 AACACCCCTCTCCACAGATTCTCATCAAGAAACACCCCGGTACCTCGGAATCCTTCGACC 4179  
QY 661 GluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSer 680  
Dbb 4180 ACCTTCAGTCCGGCAAGTTGCTTCTTTCATCACACAGTACTCCACGGGACAGGTCAGC 4239  
QY 681 ValGluIleGluTrrpGluLeuGlnLysGluAsnSerLysArgTrrpAsnProGluValGln 700  
Dbb 4240 GTGGAGATCGAGTGGGAGCTGCAGAAGGAAACAGCAAAACGCTGGAATCCCGAAATTTCAG 4299  
QY 701 TyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeu 720  
Dbb 4300 TACACTTCCAACACTACAACAAGTCTGTTAATGTGGACTTTACTGTGGACACTAATGGCGTG 4359  
QY 721 TyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736  
Dbb 4360 TATTCAGAGCCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTG 4407  
RESULT 9  
ABK89694  
ID ABK89694 standard; DNA; 4679 BP.  
XX AC ABK89694;  
XX 05-NOV-2002 (first entry)  
XX Adeno-associated virus 2 (AAV2) vector.  
XX Adeno-associated virus 2 vector; AAV2; ds; cyclic; circular; cancer;  
KW VP1 capsid; heparin-sulphate proteoglycan; vaccine; immune response;  
KW ovarian cancer.  
XX Adeno-associated virus 2.  
FH Key Location/Qualifiers  
FT CDS 2203..4410  
FT /\*tag= a  
FT /product= "Adeno-associated virus 2, VP1 capsid  
FT protein"  
FT CDS 2614..4410  
FT /\*tag= b  
FT /product= "Adeno-associated virus 2, VP2 capsid  
FT protein"  
FT /transl\_except= (pos:2614..2616, aa:Met)  
FT 2809..4410  
FT /\*tag= c  
FT /product= "Adeno-associated virus 2, VP3 capsid  
FT protein"  
XX WO200253703-A2.  
XX 11-JUL-2002.  
XX 04-JAN-2002; 2002WO-US00152.  
XX 05-JAN-2001; 2001US-260124P.  
XX (CHIL-) CHILDRENS HOSPITAL INC.  
XX Bartlett JS;  
XX WPI; 2002-583608/62.  
XX P-PSDB; AAU98974, AAU98975, AAU98976.  
XX New adeno-associated virus vector comprises a biotinylated capsid or  
PT capsid protein with an amino acid insertion in the VP1 capsid, useful  
PT as a vaccine or for transferring a therapeutic peptide to a cancer cell





Db 3940 TCTACCACTCCAGAGAGGCAACAGACAAGCAGCTACCCGACATGTCAACACACAAGGC 3999

Qy 601 AlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAla 620

Db 4000 GTTCTTCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCATCTGGGCA 4059

Qy 621 LysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeu 640

Db 4060 AAGATTCCACACACGGACGGACATTTCACCCCTCTCCCTCATGGGTGGATTCCGGACTT 4119

Qy 641 LysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAla 660

Db 4120 AAACACCTCTCTCCACAGATTCTCATCAAGAAACACCCCGGTACTGCGAATCCTTCGACC 4179

Qy 661 GluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSer 680

Db 4180 ACCTTCAGTGGGCAAGTTTGCTTCCTTCATCACACAGTACTCCACGGGACAGGTCAGC 4239

Qy 681 ValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGln 700

Db 4240 GTGGAGATCGAGTGGGAGCTGCAGAAGGAAACAGCAACCGTGGATCCCGAAATTCAG 4299

Qy 701 TyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeu 720

Db 4300 TACACTTCCAACTACAACAAGTCTGTAAATGTGGACTTTACTGTGGACACTAATGGCGTG 4359

Qy 721 TyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736

Db 4360 TATTCAGAGCCTCGCCCCATTGGCACCATACCTGACTCGTAAATCTG 4407

RESULT 10

ABV76133

ID ABV76133 standard; DNA; 4679 BP.

XX AC ABV76133;

XX DT 07-MAR-2003 (first entry)

XX DE Adeno associated virus type 2 coat protein nucleic acid.

XX KW Gene therapy; vector; hepatitis B virus; cardiovascular disease;

XX KW heart; cardiant; vasotropic; antiarrhythmic; antiarteriosclerotic;

XX KW coat protein; gene; ss.

XX OS Adeno associated virus type 2.

XX PN WO200287594-A1.

XX PD 07-NOV-2002.

XX PF 30-APR-2002; 2002WO-US13644.

XX PR 30-APR-2001; 2001US-287423P.

XX PA (REGC ) UNIV CALIFORNIA.

XX PI Chien KR, Hoshijima M;

XX DR WPI; 2003-111844/10.

XX Novel non-viral vector comprises vesicular membrane with hepatitis B

PT envelope protein with cardiac targeting sequence, and nucleotide

PT sequence for gene therapy useful for treating, e.g., heart failure,

PT arrhythmia and atherosclerosis -

XX PS Disclosure; Page 24-27; 53pp; English.

XX CC The present sequence is that of an adeno-associated virus (AAV) type

CC 2 coat protein nucleic acid. The invention provides a non-viral

CC vesicle vector for the delivery of nucleic acid to various cardiac

CC cell types. The vesicle vector contains the hepatitis B virus

CC envelope protein in which at least part of the liver targeting

CC sequence is deleted and replaced with a specific cardiac cell

CC targetting sequence. For example, the loop IV region of the AAV

CC binds to heparin sulfate proteoglycans on the surface of

CC cardiomyocytes, and can be used to target the vector to

CC cardiomyocytes. The vesicle vector can be delivered intravenously

CC or intra-arterially rather than by more invasive methods such as

CC direct cardiac injection. It can be used to deliver gene products

CC to replace or enhance expression of proteins for treatment of heart

CC failure, arrhythmia, reperfusion injury, atherosclerosis, to

CC promote angiogenesis, etc. The vesicles are highly stable and can

CC be produced in large quantities, making them ideal for gene therapy.

XX SQ Sequence 4679 BP; 1198 A; 1262 C; 1255 G; 964 T; 0 other;

Alignment Scores:

Pred. No.: 7.1e-240 Length: 4679

Score: 3402.50 Matches: 613

Percent Similarity: 90.22% Conservative: 51

Best Local Similarity: 83.29% Mismatches: 71

Query Match: 85.30% Indels: 1

DB: 25 Gaps: 1

US-09-807-802A-13 (1-736) x ABV76133 (1-4679)

Qy 1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20

Db 2203 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGACACTCTCTCTGAAGGAATAAGA 2262

Qy 21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnLysGlnAsp 40

Db 2263 CAGTGGTGGAAGCTCAAAACCTGGCCACCACCACCAAGCCCGCAGACGGCATAAGGAC 2322

Qy 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyPropheAsnGlyLeuAsp 60

Db 2323 GACAGCAGGGGTCTTGTGTCTTCTGGTACAAGTACTCGGACCTTCAACGGACTCGAC 2382

Qy 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80

Db 2383 AAGGAGAGCCGGTCAACGAGGCAGACGCCCGCGCCCTCGAGCAGCACAAGCCTACGAC 2442

Qy 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100

Db 2443 CGGCAGCTCGACAGCGGAGACAACCCGTACTCTCAAGTACAACACCACGCCGCGGAGTTT 2502

Qy 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120

Db 2503 CAGGAGCGCCTTAAAGAAGATACGTCTTTTGGGGGCAACCTCGGACGAGCAGTCTTCCAG 2562

Qy 121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro 140

Db 2563 GCGAAAAAAGAGGGTTCTTGAACCTCTGGGCCTGGTTGAGAAACCTGTTAAGACGGTCCG 2622

Qy 141 GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly 160

Db 2623 GGAATAAAGAGGCCGCTAGACACTCTCTGTGGAGCCAGACTCTCTCTCGGGAACCGGA 2682

Qy 161 LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180

Db 2683 AAGCGGGCCAGCAGCCTGCAAGAAAAAGATTGAATTTTGGTCAGACTGGAGACGCAGAC 2742

Qy 181 SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro 200

Db 2743 TCAGTACCTGACCCCCCAGCCTCTCGGACAGCCACCAGCAGCCCCCTCTGGTCTGGGAAC 2802

Qy 201 ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly 220

Db 2803 AATACGATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGGCGCCGCGGA 2862

Qy 221 ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle 240

Db 2863 GTGGTAATTCCTCGGGAATTTGGCATTCGATTCCACATGGATGGGCGCAGAGTCATC 2922

Qy 241 ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle 260

Db 2923 ACCACGACCCCGAACCTGGGCCCTGCCACCCTACAACAACCACTCTACAACAAATT 2982

Qy 261 SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp 280

Db 2983 TCCAGCCCAATCA---GGAGCCTCGAACGACAACTACTACTTTGGCTACAGCACCCCTTGG 3039

Qy 281 GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu 300

Db 3040 GGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTACCACGTCGACTGGCAAGACTC 3099

Qy 301 IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln 320

Db 3100 ATCAACAACAACCTGGGATTCGACCCCAAGAGACTCAAACTTCAAGCTCTTTAAACATTCAA 3159

Qy 321 ValLysGluValThrThrAsnAspGlyValThrIleAlaAsnAsnLeuThrSerThr 340

Db 3160 GTCAAAGAGGTCACGCGAATGACGGTACGACGCGATTGCCAATAACCTTACCAGCAGC 3219

Qy 341 ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln 360

Db 3220 GTTCAGGTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTCTCTGGCTCGGCGCATCAA 3279

Qy 361 GlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr 380

Db 3280 GGATGCCTCCCGCCGTTCCAGCAGACGCTCTTCATGCTGCGTCCACAGTATGGATACCTCACC 3339

Qy 381 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 400

Db 3340 CTGAACAACGGGAGTCAGGCGAGTAGGACGCTCTTCAATTTACTGCTGGAGTACTTTCCT 3399

Qy 401 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluValPro 420

Db 3400 TCTCAGATGCTGCGTACCGGAAACAACACTTTACCTTCAGCTACACTTTTGAGGACGTTCCCT 3459

Qy 421 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 440

Db 3460 TTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTTCATCGAC 3519

Qy 441 GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAsp 460

Db 3520 CAGTACCTGTATTACTTGACGAGAACAACAACACTCCAAGTGGAAACCACCAGCAGTCAAGG 3579

Qy 461 LeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuPro 480

Db 3580 CTTCAGTTTTCTCAGGCCGAGCGAGTGACATTCGGGACCAGTCTAGGAACCTGCTTCCT 3639

Qy 481 GlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsn 500

Db 3640 GGACCTCTGTTACCGCCAGCAGCGAGTATCAAAAGACATCTGCGGATACAAACAACAGTGAA 3699

Qy 501 PheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnPro 520

Db 3700 TACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGCTGAATCCG 3759

Qy 521 GlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyVal 540

Db 3760 GGCCCGCCATGGCAAGCCACAAGGACGATGAAGAAAAGTTTTTCTCTCAGAGCGGGTT 3819

Qy 541 MetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIle 560

Db 3820 CTCATCTTTGGGAAGCAAGGCTCAGAGAAAACAATGTGGACATTGAAAAGGTCATGATT 3879

Qy 561 ThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrVal 580

Db 3880 ACAGACGAAGAGGAAATCAGGACAACCAATCCCGTGGTACGGAGCAGTATGGTTCTGTA 3939

Qy 581 AlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGly 600

Db 3940 TCTACCAACCTCCAGAGAGGCAACAGACAGCAGCTACCGCAGATGTCAACACACAAGGC 3999

Qy 601 AlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAla 620

Db 4000 GTTCTTCCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGGCA 4059

Qy 621 LysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeu 640

Db 4060 AAGATTCACACACGACGACGACATTTTCAACCCCTCTCCCTCATGGTGGATTTCGGACTT 4119

Qy 641 LysAsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAla 660

Db 4120 AAACACCCCTCTCCACAGATTCTCATCAAGAACAACCCCGGTACCTCGAATCCTTCGACC 4179

Qy 661 GluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSer 680

Db 4180 ACCTTCAGTGGGCAAAAGTTTGTCTTCTTCATCACACAGTACTCCACGGGACAGGTCAGC 4239

Qy 681 ValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGln 700

Db 4240 GTGGAGATCGAGTGGGAGCTGCAGAGGAAACAGCAACCGCTGGAAATCCCGAAATTCAG 4299

Qy 701 TyrThrSerAsnTyrAlaLysSerAlaAsnValAsnPheThrValAspAsnAsnGlyLeu 720

Db 4300 TACACTTCCAACACTACAACNAGTCTGTTAATGTGGACTTTTACTGTGGACACTAATGGCGTG 4359

Qy 721 TyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736

Db 4360 TATTCAGAGCCTGCCCCATTGGCACCCAGATACCTGACTCGTAATCTG 4407

RESULT 11

AAH26326  
ID AAH26326 standard; DNA; 7557 BP.

XX AC AAH26326;

XX DT 02-OCT-2001 (first entry)

XX DE Adeno-associated virus packaging plasmid pfloxAAV.

XX KW AAV; vector; pfloxAAV2; packaging cell line; gene therapy; ds.

XX OS Adeno associated virus.

XX PN WO200155361-A2.

XX PD 02-AUG-2001.

XX PF 26-JAN-2001; 2001WO-US02709.

XX PR 26-JAN-2000; 2000US-0178536.

XX PA (CHIR ) CHIRON CORP.

XX PI Hardy SF;

XX DR WPI; 2001-483239/52.

XX PT Producing recombinant adeno-associated virus (rAAV) vector, by stably infecting eukaryotic host cell with rAAV vectors, helper adenovirus or herpes virus, recombinant adenovirus or herpes vectors -

PS Example 1; Page 54-57; 63pp; English.

XX CC The present sequence is that of adeno-associated virus (AAV) packaging plasmid pfloxAAV2. The following AAV packaging plasmids were prepared: pKSrepap, a non-replicating control plasmid containing rep and cap genes but no inverted terminal repeats (ITRs); pAV2 (see AAH26325), containing a wild-type AAV genome; and pfloxAAV, a version of pAV2 with 2 loxP sites inserted such they flank the rep and cap genes and thus separate the ITRs from these genes. Each was combined with pCMV GFP (see AAH26324), a plasmid with a green fluorescent protein expressing recombinant AAV (rAAV) genome, and used to transfect 293 or Cre8 cells. After 6 hr, the medium was changed and wild-type adenovirus type 5 was added.

XX CC Virus particles were harvested 3 days later. Plasmid pAV2 produced mostly AAV and a low yield of GFP vector with no effect of Cre recombinase. pfloxAAV packaged as much GFP vector as the







Db 2683 AAGCGGGCCAGCAGCCTGCAAGAAAAAGATTGAATTTTGGTCAGACTGGAGACGCAGAC 2742

Qy 181 SerValProAspProGlnProLeuGlyGluProAlaThrProAlaAlaValGlyPro 200

Db 2743 TCAGTACCTGACCCCGAGCCTCTCGGACAGCCACCAGCAGCCCTCTGGTCTGGGAAC 2802

Qy 201 ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly 220

Db 2803 AATACGATGGCTACAGGCAGTGGCGCAACCAATGGCAGACAAATACAGAGGCGCCGACGGA 2862

Qy 221 valGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle 240

Db 2863 GTGGGTAAATCTCGGGAATTTGGCAATGGCATTCACATGGATGGGCGACAGATCATC 2922

Qy 241 ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle 260

Db 2923 ACCACGACACCCGAACCTGGGCCCTGCCACCTACAACAACCACTCTACAAACAAAT 2982

Qy 261 SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp 280

Db 2983 TCCAGCCCAATCA---GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTGG 3039

Qy 281 GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu 300

Db 3040 GGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCAACACGCTGACTGGCAAGACTC 3099

Qy 301 IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln 320

Db 3100 ATCAACAACAACCTGGGATTCGACCCCAAGAGACTCAACTTCAAGCTCTTTAACAATTCAA 3159

Qy 321 valLysGluValThrThrAsnAspGlyValThrIleAlaAsnAsnLeuThrSerThr 340

Db 3160 GTCAAAGAGGTACCGCAGAATGACGGTACGACGACGATTGCCAATAAACCCTTACCAGCACG 3219

Qy 341 valGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln 360

Db 3220 GTTCAGGTGTTTACTGACTCGGAGTACCAAGTCCCGTACGTCCTCGGCTCGGCGCATCAA 3279

Qy 361 GlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr 380

Db 3280 GGATGCCTCCCGCGTTCCAGCAGACGCTCTCATGGTGCCACAGTATGGATACCTCACCC 3339

Qy 381 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 400

Db 3340 CTGAACAACGGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCTGGAGTACTTTCTCT 3399

Qy 401 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro 420

Db 3400 TCTCAGATGCTGCTACCGGAAACAACCTTACCTTCAGCTACACTTTTGGAGACGTTCTCT 3459

Qy 421 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 440

Db 3460 TTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCTCTCATCGAC 3519

Qy 441 GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAsp 460

Db 3520 CAGTACCTGTATTACTTGAGCAGAAACAACAACCTCAAGTGGAACCAACGACGAGTCAAGG 3579

Qy 461 LeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuPro 480

Db 3580 CTTTCAGTTTCTCAGCGCGGAGCGAGTGACATTCGGGACCAAGTCTAGGAAGTGGCTTCTCT 3639

Qy 481 GlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsn 500

Db 3640 GGACCCCTGTTACCGCCAGCAGCGAGTATCAAGAATCTCGGATACCAACAACAGTGA 3699

Qy 501 PheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnPro 520

Db 3700 TACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCG 3759

Qy 521 GlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyVal 540

Db 3760 GGCCCCGGCCATGGCAAGCCACAAGGACGATGAAGAAAAGTTTTTCTCTCAGAGCGGGGTT 3819

Qy 541 MetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIle 560

Db 3820 CTCATCTTTGGGAAGCAAGGCTCAGAGAAAAACAATGTGGACATTGAAAAGGTTCATGATT 3879

Qy 561 ThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrVal 580

Db 3880 ACAGACGAAGAGGAATCAGGACAAACCAATCCCCTGGTACGGAGCAGTATGGTTCGTGA 3939

Qy 581 AlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGly 600

Db 3940 TCTACCAACCTCCAGAGAGGCAACAGACAAGCAGCTACCCGAGATGTCAACACACAAGGC 3999

Qy 601 AlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAla 620

Db 4000 GTTCTTCCAGGCATGGTCTGGCAGGACAGAGATGTGTACTTCCAGGGGCCCATCTGGGCA 4059

Qy 621 LysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeu 640

Db 4060 AAGATTCCACACACGACGACGACATTTTCAACCCCTCTCCCTCATGGGTGGATTCCGACTT 4119

Qy 641 LysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAla 660

Db 4120 AAACACCTCTCTCCACAGATTCTCATCAAGAACAACCCCGGTACCTCGAATCCTTCGACC 4179

Qy 661 GluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSer 680

Db 4180 ACCTTCAGTGGCGCAAGTTTGTCTTCTTCATCACACAGTACTCCACGGGACAGGTCAGC 4239

Qy 681 ValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGln 700

Db 4240 GTGGAGATCGAGTGGGAGCTGCAGAAGGAAAACAGCAACGCTGGAATCCCGAAATTCAG 4299

Qy 701 TyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeu 720

Db 4300 TACACTTCCAACACTACAACAAGTCTGTTAATGTGGACTTTACTGTGGACACTAATATGGCGTG 4359

Qy 721 TyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736

Db 4360 TATTTCAGAGCCTCGCCCCATTGGCACCAGATACCTGACTCGTAATCTG 4407

RESULT 13

AAT09008

ID AAT09008 standard; DNA; 4680 BP.

XX AAT09008;

AC AAT09008;

XX 05-JUL-1996 (first entry)

DE Wild-type adeno-associated virus 2 genome.

XX adeno-associated virus 2; AAV2; inverted terminal repeat; viral vector;

OS DNA delivery; immunodeficiency virus protein; immunity; human; simian;

XX neurological disorder; therapy; ss.

XX Adeno-associated virus.

XX WO9534670-A2.

XX 21-DEC-1995.

XX 06-JUN-1995; 95WO-US07178.

XX 06-JUN-1994; 94US-0254358.

XX (CHIL-) CHILDRENS HOSPITAL INC.

XX Johnson PR;

XX WPI; 1996-049697/05.

XX Recombinant adeno-associated virus genome contg. protein encoding

PT DNA - flanked by inverted terminal repeats, for use in vaccines or



PT for treatment of neuro-degenerative disease

PS Example 1; Page 27-29; 44pp; English.

XX  
CC The present sequence is that of the wild-type adeno-associated virus  
CC (AAV) 2, a replication-deficient parvovirus. Cis-acting sequences  
CC directing viral DNA replication (ori), encapsidation/packaging (pkg) and  
CC host cell chromosome integration (int) are contained within the ITRs  
CC (inverted terminal repeats of 145 nucleotides). When AAV infects a  
CC human cell, the viral genome integrates into chromosome 19 resulting in  
CC latent infection of the cell. Prodn. of infectious virus does not occur  
CC unless the cell is infected with a helper virus (e.g., adenovirus or  
CC herpes- virus). AAV possesses unique features that make it attractive as  
CC a vector for delivering foreign DNA to cells. A vector including a  
CC recombinant AAV genome contg. a simian immunodeficiency virus (SIV) rev  
CC and envelop (gpl60) gene cassette was constructed from an existing plasmid  
CC designated psb201. The plasmid contains a modified wild-type AAV2  
CC genome. Specifically, two XbaI sites were added via linker addition at  
CC sequence positions 190 and 4484. These sites are internal to 191 bp ITRs  
CC which include the 145 bp ITRs of the AAV genome. The insertion of these  
CC sites allows the complete removal of the internal 4.3 kb fragment contg.  
CC the AAV rep-cap genes upon XbaI digestion of the plasmid. (See also  
CC AAT09009-10).

XX  
SQ Sequence 4680 BP; 1198 A; 1265 C; 1254 G; 963 T; 0 other;

Alignment Scores:

Pred. No.: 8.61e-237 Length: 4680  
Score: 3360.50 Matches: 609  
Percent Similarity: 89.42% Conservative: 50  
Best Local Similarity: 82.63% Mismatches: 76  
Query Match: 84.24% Indels: 2  
DB: 17 Gaps: 1

US-09-807-802A-13 (1-736) x AAT09008 (1-4680)

QY 1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20  
DB 2203 ATGGCTGCCGATGGTTATCTCCAGATTGGCTCGAGGACACTCTCTCTGAAGGAATAAGA 2262  
QY 21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp 40  
DB 2263 CAGTGGTGAAGCTCAAACCTGGCCCCACCACCACCAAGCCCGCAGAGCGGCATAAGGAC 2322  
QY 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyPropheAsnGlyLeuAsp 60  
DB 2323 GACAGCAGGGGTCTTGTCCTCTGGGTACAAGTACCTCGACCCCTTCAACGGACTCGAC 2382  
QY 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80  
DB 2383 AAGGAGAGCCGGTCAACGAGCAGACGCCCGGCCCTCGAGCACGACAAAGCCTACGAC 2442  
QY 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100  
DB 2443 CGGCAGCTCGACAGCGGAGACAAACCCGTACCTCAAGTACAACACACCGCCGAGGAGTTT 2502  
QY 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120  
DB 2503 CAGGAGCGCCTTAAGAAGATACGTCTTTGGGGCAACCTCGGACGAGCAGTCTTCCAG 2562  
QY 121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro 140  
DB 2563 GCGAAAAGAGGGTCTTTGAACCTCTGGGCCCTGGTTGAGGAACCTGTTAAGACGGCTCCG 2622  
QY 141 GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly 160  
DB 2623 GGA AAAAGAGCGCGGTAGAGCACTCTCTGTGGAGCCAGACTCTCTCTCGGGAACCGGA 2682  
QY 161 LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180  
DB 2683 AAGCGGGCCAGCAGCCTGCAAGAAAAAGATTGAATTTTGGTCAGACTGGAGACGCGAC 2742  
QY 181 SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro 200

DB 2743 TCAGTACCTGACCCCGAGCCTCTCGGACAGCCACCAGCAGCCCTCTCTGGTCTGGGAAC 2802  
QY 201 ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly 220  
DB 2803 AATACGATGGCTACAGGCAGTGGCGACCAATGGCAGACAATAACAGGGCGCCGACGGA 2862  
QY 221 ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValile 240  
DB 2863 GTGGGTAAATCTCCGGAATTTGGCATTGCGATTCCACATGGATGGGCGCAGAGTCATC 2922  
QY 241 ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle 260  
DB 2923 ACCACCAGCACCCGAACCTGGGCCCTGCCACCTACAACAACCACTCTACAACAATAAT 2982  
QY 261 SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTir 280  
DB 2983 TCCAGCCCAATCA---GGAGCCTCGAAGCACAATCACTACTTTGGCTACAGCACCCCTTG 3039  
QY 281 GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu 300  
DB 3040 GGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCACCACCGTACTGGCAAGACTC 3099  
QY 301 IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln 320  
DB 3100 ATCAACAACAACCTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAACATTCAA 3159  
QY 321 VallysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThr 340  
DB 3160 GTCAAAGAGGTCAACGAGAATGACGGTACGACGACGAGATTGCCAATAACCTTACCAGCAG 3219  
QY 341 ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln 360  
DB 3220 GTTCAGGTGTTACTGACTCGGAGTACCGAGTCCCGTACGTCTCGGCTCGGCGCATCAA 3279  
QY 361 GlyCysLeuProPropheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr 380  
DB 3280 GGATGCCTCCCGCGTTCCCGAGCAGACGCTTTCATGGTGCACAGTATGGATACCTCACC 3339  
QY 381 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 400  
DB 3340 CTGAACAACGGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCCTGGAGTACTTTCTCT 3399  
QY 401 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro 420  
DB 3400 TCTCAGATGCTGCGTACCGGAACAACATTTTACCTTCAGCTACACTTTTGGAGCGTTCTCT 3459  
QY 421 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 440  
DB 3460 TTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGCTCTCATGAATCTCTCATCGAC 3519  
QY 441 GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAsp 460  
DB 3520 CAGTACCTGTATTACTTGAGCAGACAACAACACTCCAAGTGGAAACCACCCAGCAGTCAAGG 3579  
QY 461 LeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuPro 480  
DB 3580 CTTTCAGTTTCTCAGGCCGGAGCGAGTGACATTCGGGACCAGTCTAGGAACCTGGCTTCTCT 3639  
QY 481 GlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsn 500  
DB 3640 GGACCCCTGTTTACCGCCAGCAGCGAGTATCAAAAGACATCTCGCGGATAACAACAACAGTGAA 3699  
QY 501 PheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleLeuAsnPro 520  
DB 3700 TACTCGTGAGCTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTCTGGTGAATCCG 3759  
QY 521 GlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyVal 540  
DB 3760 GGGCCCGCCATGGCAAGCCACAAGCAGATGAAGAAAAAGTTTTTTCCTCAGAGCGGGGTT 3819  
QY 541 MetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIle 560

Db 3820 CTCATCTTTGGGAAGCAAGGCTCAGAGAAAACAAATGTGAACATTGAAAAGGTCATGATT 3879

Qy 561 ThrAspGluGluGluLeuLysAlaThrAsnProValAlaThrGluArgPheGlyThrVal 580  
|||||

Db 3880 ACAGACGAAGAGAAATCGGAACACCAACATCCCGTGGCTACGAGCAGTATGGTTCTGTA 3939  
|||||

Qy 581 AlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGly 600  
::: ||| ::| |||||

Db 3940 TCTACCAACCTCCAGAGAGGCAACAGACAGCAGCTACCGCAGATGTCAACACACAGGC 3999  
|||||

Qy 601 AlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAla 620  
|||||

Db 4000 GTTCTCCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGCA 4059  
|||||

Qy 621 LysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeu 640  
|||||

Db 4060 AAGATTCCACACACGGACGGACATTTTACCCCTCTCCCTCATGGTGGATTCCGGACTT 4119  
|||||

Qy 641 LysAsnProProGlnIleLeuLysAsnThrProValProAlaAsnProProAla 660  
||:::|||||

Db 4120 AAACACCCCTCTCCACAGATTCTCATCAAGAACACACCCCGGTACCTGCGAATCCTTCGACC 4179  
|||||

Qy 661 GluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGln-Valse 680  
|||||

Db 4180 ACCTTCAGTGGGCAAGTTTGCTTCTTCATCACACAGTACTCCACGGGACACGGTCA 4239  
|||||

Qy 680 rValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValG1 700  
|||||

Db 4240 CGTGGAGATCGAGTGGGAGCTGCAGAAAGAAACAGCAACCGTGGATCCCGAAATTCA 4299  
|||||

Qy 700 nTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLe 720  
|||||

Db 4300 GTACACTTCCAACACTACAACAGTCTGTTAATCGTGGACTTACCGTGGATACATAATGGCGT 4359  
|||||

Qy 720 uTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736  
: ||:::|||||

Db 4360 GTATTTCAGAGCGCTCGCCCCCATTTGGCACGAGATACCTGACTCGTAATCTG 4408  
|||||

RESULT 14

ABX14497

ID ABX14497 standard; DNA; 4680 BP.

XX AC ABX14497;

XX DT 16-APR-2003 (first entry)

XX DE Wild-type adeno-associated virus 2 (AAV2) genomic sequence.

XX KW Recombinant adeno-associated virus; rAAV; inverted terminal repeat; ITR; DNA delivery; gene therapy; immunity; therapeutic gene delivery; human immunodeficiency virus type 1; HIV-1; neurological disorder; KW Alzheimer's disease; Parkinson's disease; Huntington's disease; KW acquired immunodeficiency syndrome; AIDS; cancer; autoimmune disease; KW multiple sclerosis; arthritis; depression; pain disorder; measles; KW seizure disorder; respiratory tract infection; mumps; hepatitis; KW neurotropic; neuroprotective; cytostatic; analgesic; anticonvulsant; KW virucide; hepatotropic; adeno-associated virus 2; AAV2; ds.

XX OS Adeno-associated virus 2.

XX PN US2002159979-A1.

XX PD 31-OCT-2002.

XX PF 15-FEB-2002; 2002US-0077294.

XX PR 06-JUN-1995; 95US-0466606.

XX PR 22-JAN-1998; 98US-0012132.

XX PR 15-APR-1999; 99US-0292703.

XX PR 18-OCT-2000; 2000US-0691604.

XX PR 06-JUN-1994; 94US-0254358.

XX PA (CHIL-) CHILDRENS HOSPITAL INC.

XX Johnson PR;

XX PI

XX WPI; 2003-209210/20.

XX PT New recombinant adeno-associated virus genome, useful for treating HIV, PT neurological disorders or cancer, comprises adeno-associated virus PT inverted terminal repeats flanking DNA sequences, and a polyadenylation PT sequence -

XX PS Example 1; Page 8-10; 19pp; English.

XX CC The present invention relates to recombinant adeno-associated virus (rAAV) genomes comprising adeno-associated virus inverted terminal repeats (ITRs) flanking DNA sequences encoding an immunodeficiency virus protein operably linked to a promoter and polyadenylation sequences. The recombinant adeno-associated virus genomes are useful for DNA delivery to cells (e.g. in gene therapy). They are particularly useful for generating immunity to human immunodeficiency virus 1 (HIV-1) and in therapeutic gene delivery for treating neurological disorders (e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease), acquired immunodeficiency syndrome (AIDS), cancer, autoimmune diseases (e.g. multiple sclerosis, arthritis), depression, pain or seizure disorders, and infections (e.g. respiratory tract infections, measles, mumps, hepatitis). The present sequence representing the wild-type genomic sequence of adeno-associated virus 2 (AAV2) is used to create a rAAV vector.

XX SQ Sequence 4680 BP; 1198 A; 1265 C; 1254 G; 963 T; 0 other;

Alignment Scores:

|                        |           |               |      |
|------------------------|-----------|---------------|------|
| Pred. No.:             | 8.61e-237 | Length:       | 4680 |
| Score:                 | 3360.50   | Matches:      | 609  |
| Percent Similarity:    | 89.42%    | Conservative: | 50   |
| Best Local Similarity: | 82.63%    | Mismatches:   | 76   |
| Query Match:           | 84.24%    | Indels:       | 2    |
| DB:                    | 25        | Gaps:         | 1    |

US-09-807-802A-13 (1-736) x ABX14497 (1-4680)

Qy 1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20  
|||||

Db 2203 ATGGCTGCCGATGGTTATCTTCAGATTGGCTCGAGGACACTCTCTCTGAAGGAATAAGA 2262  
|||||

Qy 21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnLysGlnAsp 40  
::: |||||

Db 2263 CAGTGGTGGAGCTCAAACTGGCCACCACCACCAAGCCCGCAGCGGCATAGGAC 2322  
|||||

Qy 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyPropheAsnGlyLeuAsp 60  
||| |||||

Db 2323 GACAGCAGGGGTCTTGTGCTTCTGGGTACAAGTACCTCGGACCCCTTCAACGGACTCGAC 2382  
|||||

Qy 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80  
||| |||||

Db 2383 AAGGAGAGCCGGTCAACGAGGCAGACGCCGCGGCCCTCGAGCAGACAAAGCCTACGAC 2442  
|||||

Qy 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100  
::: ||||| ::: |||||

Db 2443 CGGCAGCTCGACAGCGGAGACAACCCGTACCTCAAGTACAACACCGCGGAGGTTT 2502  
|||||

Qy 101 GlnGluArgLeuGlnGluAspThrSerPheGlyAsnLeuGlyArgAlaValPheGln 120  
|||||

Db 2503 CAGGAGCGCCTTAAGAGATACGTCTTTTGGGGCAACCTCGGACGAGCAGTCTTCCAG 2562  
|||||

Qy 121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro 140  
|||||

Db 2563 GCGAAAAAGAGGGTTCTTGAACCTCTGGCCCTGGTTGAGGAACCTGTTAAGACGGCTCCG 2622  
|||||

Qy 141 GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerGlyIleGly 160  
|||||

Db 2623 GGAAGAGAGCGCGGTAGAGCACTCTCTCTGTGGAGCCAGACTCTCTCTCGGAACCGGA 2682  
|||||

Qy 161 LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180  
|||||





PT Gene transfer method for highly efficient introduction of foreign genes  
PT to cells in humans particularly by integration specifically into AAVS1  
PT site of chromosome 19 for transfection, applicable e.g. in disease  
PT treatment -  
XX  
PS Example 1; Page 28-31; 38pp; Japanese.  
XX  
CC The present invention describes a gene transfer method for highly  
CC efficient introduction of foreign genes to cells in humans particularly  
CC by integration specifically into AAVS1 site of chromosome 19 for  
CC transfection, applicable e.g. in disease treatment. The method  
CC comprises transferring into the cells (with the use of an adenovirus  
CC vector) a nucleic acid which has a sequence provided with adeno-  
CC associated virus (AAV)-originated inverted terminal repeats (ITRs) in  
CC both sides of the target foreign gene to be transferred, a second nucleic  
CC acid which has an AAV-originated rep gene and a promoter for expressing  
CC this gene and carries a stuffer sequence inserted into it sandwiched in  
CC two recombinase-recognition sequences and located between the rep gene  
CC and promoter and expressing the Rep protein under the action of the  
CC recombinase in the cells obtained previously to integrate the target  
CC foreign gene into the chromosomal DNA. The method is for transferring a  
CC foreign gene into cells particularly in human, especially by integration  
CC specifically into AAVS1 site of chromosome 19 for transfection, which  
CC is applicable in medicine, as well as cell, genetic and embryological  
CC engineering e.g. in disease treatment. The method is highly efficient,  
CC with use of adenovirus vectors, nucleic acids and other sequences  
CC including rep genes and promoters, by expressing rep proteins to  
CC integrate target foreign gene into chromosomal DNA through action of  
CC recombinase. The present sequence represents an AAV plasmid pAV1 Avail  
CC nucleotide sequence which is used in an example from the present  
CC invention.  
XX  
SQ Sequence 4675 BP; 1198 A; 1262 C; 1251 G; 964 T; 0 other;

Alignment Scores:  
Pred. No.: 5.3e-234 Length: 4675  
Score: 3322.50 Matches: 607  
Percent Similarity: 89.02% Conservative: 50  
Best Local Similarity: 82.25% Mismatches: 76  
Query Match: 83.29% Indels: 5  
DB: 22 Gaps: 2

US-09-807-802A-13 (1-736) x AAH41481 (1-4675)

QY 1 MetAlaAlaAspGlyTyrLeuProAspTTrpLeuGluAspAsnLeuSerGluGlyIleArg 20  
DB 2203 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACACTCTCTCTGAAGGAATAAGA 2262  
QY 21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnLysGlnAsp 40  
DB 2263 CAGTGGTGAAGCTCAAAACCTGGCCACCACCAACCAAGCCCGCAGAGCGGCATAAGGAC 2322  
QY 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp 60  
DB 2323 GACAGCAGGGGTCTTGTCTTCTGGGTACAAAGTACCTCGGACCCCTTCAACGGACTCGAC 2382  
QY 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHis-AspLysAlaTyrAs 80  
DB 2383 AAGGAGAGCCGCTCAACGAGGACAGCGCGCGCCCTCGAGCACGTACAAAGCCTACGA 2442  
QY 80 pGlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPh 100  
DB 2443 CCGCAGCTCGACAGCGGAGACAACCCGTACCTCAAGTACAACCAACCGCGCGGAGTT 2502  
QY 100 eGlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheG1 120  
DB 2503 TCAGGAGCGCCTTAAAGAAGATACGTCTTTTGGGGCAACCTCGGACGACGACTCTTCCA 2562  
QY 120 nAlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPr 140  
DB 2563 GCGGAAAAAGAGGGTCTTGAACCTCTGGGCCCTGGTTGAGGAACCTGTTAAGACGGCTCC 2622  
QY 140 oGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerGlyIleG1 160

DB 2623 GGGAAAAAGAGCGCGGTAGAGCACTCTCTGTGGAGCCAGACTCTCTCTCGGAACCGG 2682  
QY 160 YLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerG1 180  
DB 2683 AAAGGGGGCCAGCAGCCTGCAAGAAAAAGATTGAATTTTGGTCAGACTGGAGACGCGA 2742  
QY 180 uSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPr 200  
DB 2743 CTCAGTACCTGACCCCGCAGCCTCTCGGACAGCCACAGCAGCCCCCTCTGTCTGGGAAC 2802  
QY 200 oThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspG1 220  
DB 2803 TAATACGATGGCTACAGGCGAGTGGCGCACCAATGCGAGACAATAACGAGGCGCGCAGCG 2862  
QY 220 yValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgVal11 240  
DB 2863 AGTGGGTAATTCTCCGGAATTTGGCATTTGGATTCCACATGGATGGGCGCAGAGTCAT 2922  
QY 240 eThrThrSerThrArgThrTrpAlaLeuProThrTrpAsnAsnHisLeuTyrLysGln11 260  
DB 2923 CACCACAGCACCCGAACTGGGCGCTGCCCCACCTTACAAACACCCTCTACAAACAAT 2982  
QY 260 eSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTr 280  
DB 2983 TTCCAGCCCAATCA---GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTG 3039  
QY 280 pGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLe 300  
DB 3040 GGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCCACCACGTGACTGGCAAGACT 3099  
QY 300 uIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleG1 320  
DB 3100 CATCAACAACAACACTGGGGATTCGACCCCAAGAGACTCAAGCTCTTTAACAATTCA 3159  
QY 320 nValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerTh 340  
DB 3160 AGTCAAGAGGTCACGCGAGAATGACGGTACGACGAGCATTGCCAATAACCTTACCAGCAC 3219  
QY 340 rValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisG1 360  
DB 3220 GGTTCAGGTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTCTCGGCTCGGCGCATCA 3279  
QY 360 nGlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuTh 380  
DB 3280 AGGATGCCTCCCGCGGTTCCCGACGAGACGTCTTTCATGTGTCACAGTATGGATACCTCAC 3339  
QY 380 rLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePr 400  
DB 3340 CCTGAACAACGGGAGTCAGGAGTAGGACGCTCTTTCATTTTACTGCTGGAGTACTTTCC 3399  
QY 400 oSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPr 420  
DB 3400 TTCTCAGATGCTGCGTACCGGAAACAACCTTTACCTTCAGCTACACTTTTGAGGACGTTCC 3459  
QY 420 oPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAs 440  
DB 3460 TTTCCACAGCAGCTACGCTCAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGA 3519  
QY 440 pGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAs 460  
DB 3520 CCAGTACCTGTATTACTTTGAGCAGAACAAACACTCCCAAGTGGAAACCCACCGCAGTCAAG 3579  
QY 460 pLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuPr 480  
DB 3580 GCTTCAGTTTCTCAGGCCGAGGAGTGACATTCGGGACCAAGTCTAGGAACCTGGCTTCC 3639  
QY 480 oGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnSerAs 500  
DB 3640 TGGACCTGTATCCGCCAGCAGGAGTATCAAAGACATCTCGCGGATAACAACAACAGTGA 3699  
QY 500 nPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnPr 520

|    |      |  |      |
|----|------|--|------|
| Db | 3700 | ATACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCC   | 3759 |
| Qy | 520  | oGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyVa   | 540  |
| Db | 3760 | G-----GCCATGGCAAGCCACAAGGACGATGAAGAAAGTTTTTTCCTCAGAGCGGGGT     | 3813 |
| Qy | 540  | lMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIl   | 560  |
| Db | 3814 | TCTCATCTTTGGGAAGCAAGGCTCAGAGAAACAAATGTGAACATTGAAAAGGTCATGAT    | 3873 |
| Qy | 560  | eThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrVa      | 580  |
| Db | 3874 | TACAGACGAAGAGGAAATCGGAACAACCAATCCCGTGGCTACGGAGCAGATGTTCTGT     | 3933 |
| Qy | 580  | lAlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGl      | 600  |
| Db | 3934 | ATCTACCAACCTCCAGAGAGGCAACAGACAGCAGCTACCGCAGATGTCAACACACACAAGG  | 3993 |
| Qy | 600  | yAlaLeuProGlyMetValTyrGlnAspArgAspValTyrLeuGlnGlyProIleTyrAl   | 620  |
| Db | 3994 | CGTTCTTCCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGC   | 4053 |
| Qy | 620  | aLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLe   | 640  |
| Db | 4054 | AAAGATTCCACACAGGACGGACATTTTCAACCCTCTCCCTCATGGGTGGATTCTGGACT    | 4113 |
| Qy | 640  | uLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAl      | 660  |
| Db | 4114 | TAAACACCCTCTCCACAGATTCTCATCAAGAAACACCCCGGTACCTGCGAATCCTTCGAC   | 4173 |
| Qy | 660  | aGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGln-Vals   | 680  |
| Db | 4174 | CACCTTCAGTGGCGCAAGTTTGCTTCTTCATCACACAGTACTCCACGGGACACGGTCA     | 4233 |
| Qy | 680  | erValGluIleGluTyrGluLeuGlnLysGluAsnSerLysArgTyrAsnProGluValG   | 700  |
| Db | 4234 | CGGTGGAGATCGAGTGGGAGCTGCAGAGGAAACAGCAACGCTGGAATCCCGAAATTC      | 4293 |
| Qy | 700  | lnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyL   | 720  |
| Db | 4294 | AGTACACTTCCAACACTACAACAAGTCTGTTAATCGTGGACTTACCGTGGATACTAATGGCG | 4353 |
| Qy | 720  | euTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu             | 736  |
| Db | 4354 | TGTATTTCAGAGCCTCGCCCCCATTTGGCACCAGATACTGACTCGTAATCTG           | 4403 |

Search completed: February 14, 2004, 17:19:59  
Job time : 683.843 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 14, 2004, 16:11:58 ; Search time 108.687 Seconds  
(without alignments)  
2988.931 Million cell updates/sec

Title: US-09-807-802A-13  
Perfect score: 3989  
Sequence: 1 MAADGYLPDWLEDNLSEGR.....NNGLYTEPRPIGTRYLTRPL 736

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO spool/US09807802/runat\_11022004\_175609\_15955/app\_query.fasta\_1.2389  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09807802@cgn\_1\_1\_131 @runat\_11022004\_175609\_15955 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCKS=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description      |
|------------|--------|-------------|--------|-------|------------------|
| 1          | 3439   | 86.2        | 8179   | 4     | US-09-438-268-5  |
| 2          | 3402.5 | 85.3        | 4072   | 4     | US-09-770-315-4  |
| 3          | 3402.5 | 85.3        | 7557   | 4     | US-09-770-315-3  |
| 4          | 3402.5 | 85.3        | 8698   | 4     | US-09-770-315-2  |
| 5          | 3360.5 | 84.2        | 4680   | 1     | US-08-254-358-1  |
| 6          | 3360.5 | 84.2        | 4680   | 1     | US-08-475-391-1  |
| 7          | 3360.5 | 84.2        | 4680   | 2     | US-08-709-609-1  |
| 8          | 3360.5 | 84.2        | 4680   | 5     | PCT-US95-07178-1 |
| 9          | 3038.5 | 76.2        | 8151   | 4     | US-09-438-268-2  |
| 10         | 2486.5 | 62.3        | 2208   | 4     | US-09-532-594B-5 |
| c 11       | 2486.5 | 62.3        | 7214   | 4     | US-09-438-268-1  |
| 12         | 2481.5 | 62.2        | 4767   | 4     | US-09-532-594B-1 |

|      |        |      |         |   |                      |                    |
|------|--------|------|---------|---|----------------------|--------------------|
| 13   | 1830.5 | 45.9 | 1800    | 4 | US-09-532-594B-17    | Sequence 17, Appl  |
| 14   | 1690.5 | 42.4 | 1617    | 4 | US-09-532-594B-19    | Sequence 19, Appl  |
| 15   | 1410   | 35.3 | 2271    | 4 | US-09-438-268-3      | Sequence 3, Appl   |
| 16   | 439    | 11.0 | 5049    | 1 | US-08-336-345-1      | Sequence 1, Appl   |
| 17   | 439    | 11.0 | 5049    | 1 | US-08-336-345-2      | Sequence 2, Appl   |
| 18   | 439    | 11.0 | 5049    | 2 | US-08-647-655-1      | Sequence 1, Appl   |
| 19   | 439    | 11.0 | 5049    | 2 | US-08-647-655-2      | Sequence 2, Appl   |
| 20   | 263    | 6.6  | 2062    | 6 | 5223424-3            | Patent No. 5223424 |
| 21   | 247    | 6.2  | 1740    | 1 | US-07-969-213-1      | Sequence 1, Appl   |
| 22   | 220    | 5.5  | 1820    | 6 | 5223424-12           | Patent No. 5223424 |
| 23   | 203    | 5.1  | 2254    | 3 | US-08-552-369-1      | Sequence 1, Appl   |
| 24   | 199.5  | 5.0  | 5910    | 1 | US-08-195-814-1      | Sequence 1, Appl   |
| 25   | 192    | 4.8  | 1752    | 3 | US-09-022-949-1      | Sequence 1, Appl   |
| 26   | 184    | 4.6  | 1755    | 2 | US-08-317-785-1      | Sequence 1, Appl   |
| 27   | 132    | 3.3  | 3906    | 4 | US-09-996-243-83     | Sequence 83, Appl  |
| c 28 | 131    | 3.3  | 4188    | 4 | US-09-252-991A-13774 | Sequence 13774, A  |
| 29   | 131    | 3.3  | 8211    | 4 | US-09-252-991A-13656 | Sequence 13656, A  |
| c 30 | 130    | 3.3  | 1896    | 1 | US-08-198-446B-8     | Sequence 8, Appl   |
| c 31 | 130    | 3.3  | 1896    | 2 | US-08-870-693-8      | Sequence 8, Appl   |
| 32   | 127.5  | 3.2  | 4563    | 4 | US-09-252-991A-930   | Sequence 930, App  |
| c 33 | 126.5  | 3.2  | 8438    | 1 | US-07-945-283-1      | Sequence 1, Appl   |
| c 34 | 126    | 3.2  | 4403765 | 3 | US-09-103-840A-2     | Sequence 2, Appl   |
| c 35 | 126    | 3.2  | 4411529 | 3 | US-09-103-840A-1     | Sequence 1, Appl   |
| c 36 | 125    | 3.1  | 4214    | 4 | US-09-122-135-1      | Sequence 1, Appl   |
| 37   | 123.5  | 3.1  | 3451    | 4 | US-09-811-286-1      | Sequence 1, Appl   |
| 38   | 123    | 3.1  | 2997    | 4 | US-09-252-991A-11853 | Sequence 11853, A  |
| 39   | 122    | 3.1  | 2817    | 4 | US-09-620-312D-1085  | Sequence 1085, Ap  |
| 40   | 121    | 3.0  | 2305    | 4 | US-09-016-434-1282   | Sequence 1282, Ap  |
| 41   | 119.5  | 3.0  | 4319    | 3 | US-08-296-791-1      | Sequence 1, Appl   |
| 42   | 119.5  | 3.0  | 4319    | 5 | PCT-US95-10661A-1    | Sequence 1, Appl   |
| 43   | 119    | 3.0  | 3288    | 4 | US-09-107-532A-201   | Sequence 201, App  |
| 44   | 118.5  | 3.0  | 3546    | 3 | US-08-872-757-3      | Sequence 3, Appl   |
| 45   | 118.5  | 3.0  | 3546    | 4 | US-09-850-048A-3     | Sequence 3, Appl   |

ALIGNMENTS

RESULT 1  
US-09-438-268-5  
; Sequence 5, Application US/09438268  
; Patent No. 6491907  
; GENERAL INFORMATION:  
; APPLICANT: Rabinowitz, Joseph E.  
; APPLICANT: Samulewski, Richard J  
; APPLICANT: Xiao, Weidong  
; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING  
; TITLE OF INVENTION: THE SAME  
; FILE REFERENCE: 5470-186  
; CURRENT APPLICATION NUMBER: US/09/438,268  
; CURRENT FILING DATE: 1999-11-10  
; EARLIER APPLICATION NUMBER: 60/107,840  
; EARLIER FILING DATE: 1998-11-10  
; EARLIER APPLICATION NUMBER: 60/123,651  
; EARLIER FILING DATE: 1999-03-10  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 8179  
; TYPE: DNA  
; ORGANISM: Virus  
US-09-438-268-5

Alignment Scores:  
Pred. No.: 0  
Score: 3439.00  
Percent Similarity: 90.50%  
Best Local Similarity: 84.94%  
Query Match: 86.21%  
DB: 4  
Length: 8179  
Matches: 626  
Conservative: 41  
Mismatches: 68  
Indels: 2  
Gaps: 2

US-09-807-802A-13 (1-736) x US-09-438-268-5 (1-8179)

.Qy 1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20



Db 2133 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACACTCTCTCTGAAGGAATAAGA 2192  
Qy 21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp 40  
Db 2193 CAGTGGTGAAGCTCAAACCTGGCCCAACCACCAAGAGCCCGCAGAGCGGCATAGGAC 2252  
Qy 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp 60  
Db 2253 GACAGAGGGGTCTTGTGCTTCTGGGTACAAGTACCTCGGACCCCTTCAACGGGACTCGAC 2312  
Qy 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80  
Db 2313 AAGGAGAGCCGGTCAACGAGGACGACGCCGGCCCTCGAGCACGACAAGCCCTACGAC 2372  
Qy 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100  
Db 2373 CGGCAGCTCGACAGCGGAGACAACCCGTACCTCAAGTACAACACGACGCCGCGGAGTTT 2432  
Qy 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120  
Db 2433 CAGGAGCGCTTAAAGAAGATACGTCTTTTGGGGCAACCTCGGACGAGCAGTCTTCCAG 2492  
Qy 121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGluGlyAlaLysThrAlaPro 140  
Db 2493 GCGAAAAGAGGGTCTTTGAACCTCTGGGCCTGTTTGAGGAACCTGTTAAGACGGCTCCG 2552  
Qy 141 GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly 160  
Db 2553 GGAAAAAAGAGCCGGTAGAGCACTCTCTGTGGAGCCAGACTCTCTCTCGGGAAACCGGA 2612  
Qy 161 LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180  
Db 2613 AAGGGGGCAGCAGCCTGCAAGAAAAAGATTGAAATTTTGGTCAGACTGGAGACGCGAC 2672  
Qy 181 SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro 200  
Db 2673 TCAGTACCTGACCCCGCAGCCTCTCGGACAGCCACCAGCAGCCCTCTGGTCTGGGAAC 2732  
Qy 201 ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGlnGlyAlaAspGly 220  
Db 2733 AATACGATGGTTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGGCGCGCACGGA 2792  
Qy 221 ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle 240  
Db 2793 GTGGGTAAATCTCCGGAATTTGGCATTTGGCATTTCCACATGGATGGCGCAGAGTCAIC 2852  
Qy 241 ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle 260  
Db 2853 ACCACGACACCCGAACCTGGGCCCTGCCACCTACAACAACCACTCTACAAACAAT 2912  
Qy 261 SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp 280  
Db 2913 TCCAGCCAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTGG 2969  
Qy 281 GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu 300  
Db 2970 GGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCCACCAGTGAATGGCAAAGACTC 3029  
Qy 301 IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln 320  
Db 3030 ATCAACAACAACCTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAACAATTCAA 3089  
Qy 321 ValLysGluValThrThrAsnAspGlyValThrIleAlaAsnAsnLeuThrSerThr 340  
Db 3090 GTCAAGAGAGGTACGCAGAAATGACGGTACGACGACGATTGCCAATAAACCTTACCAGCACG 3149  
Qy 341 ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln 360  
Db 3150 GTTCAGGTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTGCTCGGTGCGGCACCAA 3209  
Qy 361 GlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr 380

Db 3210 GGCTGTCTCCCGCCGTTTCCAGCGGACGCTTCTTCATGGTCCCTCAGTATGGATACCTCACC 3269  
Qy 381 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 400  
Db 3270 CTGAACAACGGAAAGTCAAGCGGTGGGACGCTCATCTTTTACTGCTCGGAGTACTTCCCT 3329  
Qy 401 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro 420  
Db 3330 TCGCAGATGCTAAGGACTGGAAATAACTTCCAATTCAGCTATACCTTCGAGGATGTACCT 3389  
Qy 421 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 440  
Db 3390 TTTTCACAGCAGCTACGCTCACAGCCAGAGTTTGGATCGCTTGATGAATCCTCTTATTGAT 3449  
Qy 441 GlnTyrLeuTyrTyrLeuAsnArgThrGln--AsnGlnSerGlySerAlaGlnAsnLys 459  
Db 3450 CAGTATCTGTACTACTGAACAGAACGCAAGGAACAACCTCTGGAACAACAACAATCA 3509  
Qy 460 AspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeu 479  
Db 3510 CGGCTGCTTTTATAGCCAGGCTGGGCCCTCAGTCTATGTCTTTCAGGCCAGAAAATGGCTA 3569  
Qy 480 ProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnSer 499  
Db 3570 CCTGGGCCCTGTACTCCGGCAACAGAGACTTTCANAAGACTGCTAAACGACACAACAACAGT 3629  
Qy 500 AsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsn 519  
Db 3630 AACTTTCTTGGACAGCGGCCAGCAAAATATCATCTCAATGGCCGCGACTCGCTGGTGAAT 3689  
Qy 520 ProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGly 539  
Db 3690 CCAGGACCAGCTATGGCCAGTCAACAAGGACGATGAAGAAAAATTTTCCCTATGCACGGC 3749  
Qy 540 ValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMet 559  
Db 3750 AATCTAATATTTGGCAAGAAGGGACAACGCGCAAGTAACGCAGAATTAGATAATGTAAATG 3809  
Qy 560 IleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThr 579  
Db 3810 ATTACGGATGAAGAAGAGATTCTGTACCACCAATCTCTGTGGCAACAGAGCAGTATGGAAC 3869  
Qy 580 ValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMet 599  
Db 3870 GTGGCAAAATACTTGCAGAGCTCAAATACAGCTCCACGACTGGAACTGTCAATCATCAG 3929  
Qy 600 GlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrp 619  
Db 3930 GGGGCCCTTACCTGGCATGGTGTGGCAAGATCGTGACGTGTACCTTCAAGGACCTATCTGG 3989  
Qy 620 AlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGly 639  
Db 3990 GCAAGATTCCTCACAGGATGGACACTTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTGGA 4049  
Qy 640 LeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProPro 659  
Db 4050 CTGAAACATCCGCCTCTCAAATCATGATCAAAAATACTCCGGTACCTCGGAATCCTTCG 4109  
Qy 660 AlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnVal 679  
Db 4110 ACCACCTTCAGTGGCGCAAGTTTGTCTCTTCATCACACAGTACTCCACGGGACAGGTC 4169  
Qy 680 SerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluVal 699  
Db 4170 ACGGTGGAGATCGAGTGGAGCTGCAGAAGGAAACAGCAAAACGCTGGAATCCCGAAAT 4229  
Qy 700 GlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGly 719  
Db 4230 CAGTACACTTCCAACTACAACAAGTCTGTTAATCGTGGACTTACCGTGGATACTAATGSC 4289  
Qy 720 LeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736  
Db 4290 GTGTATTTCAGAGCCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTG 4340

RESULT 2  
US-09-770-315-4  
; Sequence 4, Application US/09770315  
; Patent No. 6429001  
; GENERAL INFORMATION:  
; APPLICANT: Chiron Corporation  
; TITLE OF INVENTION: Recombinant AAV Packaging Systems  
; FILE REFERENCE: 20263-501  
; CURRENT APPLICATION NUMBER: US/09/770,315  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: US 60/178,536  
; PRIOR FILING DATE: 2000-01-26  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 4072  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: recombinant DNA  
US-09-770-315-4

Alignment Scores:  
Pred. No.: 0 Length: 4072  
Score: 3402.50 Matches: 613  
Percent Similarity: 90.22% Conservative: 51  
Best Local Similarity: 83.29% Mismatches: 71  
Query Match: 85.30% Indels: 1  
DB: 4 Gaps: 1

US-09-807-802A-13 (1-736) x US-09-770-315-4 (1-4072)

|    |      |   |      |
|----|------|---|------|
| QY | 1    | MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg  | 20   |
| Db | 1484 | ATGGCTGCCGATGGTTATCTTCAGATTGGCTCGAGGACACTCTCTCTGAAGGAATAAGA   | 1543 |
| QY | 21   | GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnLysGlnAsp     | 40   |
| Db | 1544 | CAGTGTGGGAAGCTCAAACTGGCCACCACCACCAAGCCGCGAGCGGCATTAAGGAC      | 1603 |
| QY | 41   | AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp  | 60   |
| Db | 1604 | GACAGCAGGGGTCTTGTCCTCTGGGTACAAGTACCTCGGACCCCTTCAACGGACTCGAC   | 1663 |
| QY | 61   | LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaIleAsp     | 80   |
| Db | 1664 | AAGGGAGCCGGTCAACGAGGCAGACGCCCGCCCTCGAGCAGCAAAAGCCTTACGAC      | 1723 |
| QY | 81   | GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe  | 100  |
| Db | 1724 | CGGACGCTCGACAGCGGAGACAAACCGTACCTCAAGTACAACACCGCCGACGCGGAGTTT  | 1783 |
| QY | 101  | GlnGluArgLeuGlnGluAspThrSerPheGlyAsnLeuGlyArgAlaValPheGln     | 120  |
| Db | 1784 | CAGGAGCGCCTTAAAGAAGATACGTCTTTTGGGGCAACCTCGGACGAGCAGTCTTCCAG   | 1843 |
| QY | 121  | AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro     | 140  |
| Db | 1844 | GCGAAAAAGAGGGTTCTTGAACCTCTGGGCCCTGGTTGAGGAACCTGTTAAGACGGCTCCG | 1903 |
| QY | 141  | GlyLysLysArgProValGlnGlnSerProGlnGluProAspSerSerSerGlyIleGly  | 160  |
| Db | 1904 | GGAAAAAGAGGCCGGTAGACACTCTCTGTGGAGCCAGACTCCTCTCGGGAAACCGGA     | 1963 |
| QY | 161  | LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu  | 180  |
| Db | 1964 | AAGCGGGCCAGCAGCCTGCAAGAAAAAGATTGAATTTTGGTCAGACTGGAGACGCAGAC   | 2023 |
| QY | 181  | SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro  | 200  |
| Db | 2024 | TCAGTACCTGACCCCCCAGCCTCTCGGACAGCCACCAGCAGCCCTCTGGTCTGGGAAC    | 2083 |

|    |      |   |      |
|----|------|---|------|
| QY | 201  | ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly       | 220  |
| Db | 2084 | AATACGATGGCTACAGGCAGTGGCGCAACCAATGGCAGACAATAACGAGGCGCCGACGGA    | 2143 |
| QY | 221  | ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle    | 240  |
| Db | 2144 | GTGGGTAATTCCTCGGGAATTTGGCATTTCCGATTCCACATGGATGGGCGCAGAGTCATC    | 2203 |
| QY | 241  | ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle    | 260  |
| Db | 2204 | ACCACCAGCACCCGAACCTGGGCCCTGCCACCTTACCAACACCACCTCTACAAACAAATT    | 2263 |
| QY | 261  | SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp    | 280  |
| Db | 2264 | TCCAGCCAATCA--GGAGCCTCGAACGACAATCACAATCTTTGGCTACAGCACCCCTTGG    | 2320 |
| QY | 281  | GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu    | 300  |
| Db | 2321 | GGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCCACCAGTACTGGCAAAAGACTC    | 2380 |
| QY | 301  | IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln    | 320  |
| Db | 2381 | ATCAACAACAACCTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAACATTCAA  | 2440 |
| QY | 321  | ValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThr    | 340  |
| Db | 2441 | GTCAAAGAGGTACGCAGAAATGACGGTACGACGACGATTGGCCAATAAACCCTTACCAGCACG | 2500 |
| QY | 341  | ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln    | 360  |
| Db | 2501 | GTTCAGGTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTCTCTCGGCTCGGCGCATCAA   | 2560 |
| QY | 361  | GlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr       | 380  |
| Db | 2561 | GGATGCCTCCCGCGTTCCAGCAGACGCTCTTCATGGTGGCACAGATATGGATACCTCACC    | 2620 |
| QY | 381  | LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro    | 400  |
| Db | 2621 | CTGAACAACCGGGAGTCAGGCAGTAGGACGCTCTTCACTTTACTGCTGGAGTACTTTCCT    | 2680 |
| QY | 401  | SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro    | 420  |
| Db | 2681 | TCTCAGATGCTGCTACCGGAAACAACCTTTACCTTCAGCTACACCTTTTGAGGACGTTCTCT  | 2740 |
| QY | 421  | PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp    | 440  |
| Db | 2741 | TTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGAC    | 2800 |
| QY | 441  | GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAsp    | 460  |
| Db | 2801 | CAGTACCTGTATTACTTGAGCAGAGAAACAACCTCCAAGTGGAAACCACCACGAGTCAAGG   | 2860 |
| QY | 461  | LeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuPro    | 480  |
| Db | 2861 | CTTCAGTTTCTCAGGCCGAGCGAGTGACATTCGGGACCAGTCTAGGAACCTGGCTTCCT     | 2920 |
| QY | 481  | GlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsn    | 500  |
| Db | 2921 | GGACCTGTATTACCGCCAGCAGCGAGTATCAAAGACATCTCGGATAACAACAACAGTGAA    | 2980 |
| QY | 501  | PheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnPro    | 520  |
| Db | 2981 | TACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCG    | 3040 |
| QY | 521  | GlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyVal    | 540  |
| Db | 3041 | GGCCCCGCCCATGGCAAGCCCAAGGACGATGAAGAAAAAGTTTTCCTCAGAGCGGGGTT     | 3100 |
| QY | 541  | MetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIle    | 560  |
| Db | 3101 | CTCATCTTTGGGAAGCAAGGCTCAGAGAAAAACAATGTGGACATTGMAAAGGTCAATGATT   | 3160 |
| QY | 561  | ThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrVal    | 580  |





```
Qy 401 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro 420
Db 3428 TCTCAGATGCTGCGTACCGGAAACAACCTTTACCTTCAGCTACACTTTTGAGGACGTTCTT 3487
Qy 421 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProIleLeuAsp 440
Db 3488 TTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACGCTCTCATGAATCCTCTCATCGAC 3547
Qy 441 GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAsp 460
Db 3548 CAGTACCTGTATTACTTGAGCAGAACAAACACTCCAAAGTGGAAACCACCACGCGCTCAAGG 3607
Qy 461 LeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuPro 480
Db 3608 CTTCAGTTTCTCAGGCCGAGCGAGTGACATTCGGGACCAAGTCTAGGAACCTGGCTTCTT 3667
Qy 481 GlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsn 500
Db 3668 GGACCCTGTTACCGCCAGCAGCGAGTATCAAGACATCTCGCGGATAACAACAACAGTGAA 3727
Qy 501 PheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleLeuAsnPro 520
Db 3728 TACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCG 3787
Qy 521 GlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyVal 540
Db 3788 GGCCCGCCCATGGCAAGCCACAAGACGATGAAGAAAAGTTTTCCTCAGAGCGGGGTT 3847
Qy 541 MetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIle 560
Db 3848 CTTCATCTTTGGGAAGCAAGGCTCAGAGAAAAACAAATGTGGACATTGAAAGGTCATGATT 3907
Qy 561 ThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrVal 580
Db 3908 ACAGACGAAGAGGAAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTATGTTCTGTA 3967
Qy 581 AlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGly 600
Db 3968 TCTACCAACCTCCAGAGAGGCAACAGACAGCTACCGCAGATGTCAACACACACAGGC 4027
Qy 601 AlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAla 620
Db 4028 GTTCTTCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGCA 4087
Qy 621 LysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeu 640
Db 4088 AAGATTCCACACACGGACGGACATTTTCACCCCTCTCCCTCATGGTGGATTTCGGAATT 4147
Qy 641 LysAsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAla 660
Db 4148 AAACACCTCCTCCACAGATTCTCATCAAGAAACACCCCGGTACCTGCGAATCCTTCGACC 4207
Qy 661 GluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSer 680
Db 4208 ACCTTCAGTGGCGCAAGTTTGCTTCTTCATCACACAGTACTCCACGGGACAGGTCAGC 4267
Qy 681 ValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGln 700
Db 4268 GTGGAGATCGAGTGGGAGCTGCAGAAGGAAAACAGCAAACGCTGGAATCCCGAAATTCAG 4327
Qy 701 TyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeu 720
Db 4328 TACACTCCAACTACAACAAGTCTGTTAATGTGGACTTTTACTGTGGACACTAATGGCGTG 4387
Qy 721 TyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736
Db 4388 TATTGAGGCTCGCCCCATTGGCACCAAGATACCTGACTCGTAATCTG 4435
```

RESULT 4

```
US-09-770-315-2
; Sequence 2, Application US/09770315
; Patent No. 6429001
```

```
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; TITLE OF INVENTION: Recombinant AAV Packaging Systems
; FILE REFERENCE: 20263-501
; CURRENT APPLICATION NUMBER: US/09/770,315
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,536
; PRIOR FILING DATE: 2000-01-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 8698
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: recombinant DNA
US-09-770-315-2

Alignment Scores:
Pred. No.: 0 Length: 8698
Score: 3402.50 Matches: 613
Percent Similarity: 90.22% Conservative: 51
Best Local Similarity: 83.29% Mismatches: 71
Query Match: 85.30% Indels: 1
DB: 4 Gaps: 1

US-09-807-802A-13 (1-736) x US-09-770-315-2 (1-8698)

Qy 1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20
Db 2203 ATGGCTGCCGATGTTATCTTCCAGATTGGCTCGAGGACACTCTCTCTGAAGGAATAAGA 2262
Qy 21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp 40
Db 2263 CAGTGGTGAAGCTCAAACCTGGCCCAACCCACCACCAAGCCCGCAGAGCGGCATAAGGAC 2322
Qy 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp 60
Db 2323 GACAGCAGGGGCTTGTGCTTCTCTGGGTACAAGTACCTCGGACCCCTTCAACGGACTCGAC 2382
Qy 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80
Db 2383 AAGGAGAGCCCGTCAACGAGGCAGACGCCGCCCTTCGAGCACGACAAAGCCTACGAC 2442
Qy 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100
Db 2443 CGGCAGCTCGACAGCGGAGACAACCCGTTACCTCAAGTACAACCAACGCCGCGAGGTTT 2502
Qy 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120
Db 2503 CAGGAGCGCCTTAAAGAAGATACGTCTTTTGGGGGCAACCTCGGACGAGCAGTCTTCCAG 2562
Qy 121 AlaLysLysArgValLeuGluProLeuGlyValLeuGluGlyAlaLysThrAlaPro 140
Db 2563 GCGAAAAGAGGGTCTTTGAACCTCTGGGCCCTGTTGAGGAACCTGTTAAGACGCTCCG 2622
Qy 141 GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly 160
Db 2623 GGAAAAAAGAGGCCGCTAGAGCACTCTCCTGTGGAGCCAGACTCCTCCTCGGAACCCGA 2682
Qy 161 LysThrGlyGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180
Db 2683 AAGCGGGCCAGCAGCCTGCAAGAAAAAGATTGAATTTTGGTCAGACTGGAGACGACAGAC 2742
Qy 181 SerValProAspProGlnProLeuGlyGluProAlaThrProAlaAlaValGlyPro 200
Db 2743 TCAGTACCTGACCCCGAGCCTCTCGGACAGCCACCAAGCCAGCAGCCCCCTCTGTCTGGGA 2802
Qy 201 ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly 220
Db 2803 AATACGATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGGCGCCGACGGA 2862
Qy 221 ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle 240
```

Db 2863 GTGGTAATTCCTCGGGAATTGGCAATTCGATTCACATGGATGGGCACAGAGTCATC 2922  
QY ThrThrSerThrArgThrTrrAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle 260  
Db 2923 ACCACGACACCCGAACCTGGGCCCTGCCACCTACAAACACACCTCTACAAACAAAT 2982  
QY SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp 280  
Db 2983 TCCAGCCCAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTGG 3039  
QY GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu 300  
Db 3040 GGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCACCACTGACTGGCAAGACTC 3099  
QY IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPhelLysLeuPheAsnIleGln 320  
Db 3100 ATCAACAACAACCTGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAACAATCAA 3159  
QY ValLysGluValThrThrAsnAspGlyValThrIleAlaAsnAsnLeuThrSerThr 340  
Db 3160 GTCAAGAGGTACGCGAGAATGACGGTACGACGACGATTGCCAATAACCTTACCAGCACG 3219  
QY ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln 360  
Db 3220 GTTCAGGTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTCCTCGGCTCGGCGCATCAA 3279  
QY GlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr 380  
Db 3280 GGATGCCTCCCGCGTTCCAGCAGACGCTTTCATGGTGCCACAGTATGGATACCTCACC 3339  
QY LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 400  
Db 3340 CTGAACAACGGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCTGGAGTACTTTTCT 3399  
QY SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro 420  
Db 3400 TCTCAGATGCTGCGTACCGGAAACAACCTTACCTTCAGCTACACTTTTGAGGACGTTTCT 3459  
QY PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 440  
Db 3460 TTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGCTCTCATGAATCCTCTCATCGAC 3519  
QY GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAsp 460  
Db 3520 CAGTACCTGTATTACTTGAGCAGACAACAACACTCCAAGTGGAAACCAACACGAGTCAAGG 3579  
QY LeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuPro 480  
Db 3580 CTTCAGTCTTCTCAGGCCGGAGCGAGTGACATTCGGGACCACTCTAGGAACCTGGCTTCT 3639  
QY GlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsn 500  
Db 3640 GGACCTGTATTACCCGACGAGGAGTATCAAGACATCTCGCGGATAACAACAACAGTGAA 3699  
QY PheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnPro 520  
Db 3700 TACTCGTGGACTGGAGCTACCAAGTACCCTCAATGGCAGAGACTCTCTGGTGAATCCG 3759  
QY GlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyVal 540  
Db 3760 GGCCCGGCCCATGGCAAGCCCAAGGACGATGAAGAAAGTTTTTCTCAGAGCGGGTT 3819  
QY MetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIle 560  
Db 3820 CTCATCTTTGGGAAGCAAGGCTCAGAGAAACAAATGTGGACATTGAAGAGGTCTATGATT 3879  
QY ThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrVal 580  
Db 3880 ACAGACGAAGAGGAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTATGGTTCTGTA 3939  
QY AlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGly 600

Db 3940 TCTACCAACCTCCAGAGAGGCAACAGACAGACAGCTACCGCAGATGTCAACACACAAGGC 3999  
QY AlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAla 620  
Db 4000 GTTCTTCCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGCA 4059  
QY LysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeu 640  
Db 4060 AAGATTCCACACACGACGGACGACATTTTTCACCCCTCTCCCTCATGGGTGGATTGGACTT 4119  
QY LysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAla 660  
Db 4120 AAACACCTCTCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAATCCTTCGACC 4179  
QY GluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSer 680  
Db 4180 ACCTTCAGTGGGCAAAAGTTTGTCTTCTTCATCACACAGTACTCCACGGGACAGGTGAGC 4239  
QY ValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGln 700  
Db 4240 GTGGAGATCGAGTGGAGCTGCAGAGGAAACAGCAACGCTGGAATCCCGAAATTCAG 4299  
QY TyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeu 720  
Db 4300 TACACTTCCAACACTACAACAAGTCTGTTAATGTGGACTTACTGTGGACACTAATGGCGTG 4359  
QY TyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736  
Db 4360 TATTCAGAGCCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTG 4407

RESULT 5

US-08-254-358-1  
; Sequence 1, Application US/08254358  
; Patent No. 5658785  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Philip R.  
; TITLE OF INVENTION: Adeno-Associated Virus Materials and  
; TITLE OF INVENTION: Methods  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 S. Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/254,358  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5658785and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 31975  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4680 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-254-358-1

Alignment Scores:

|  |           |   |      |
|--|-----------|---|------|
| Pred. No.:   | 9.71e-315 | Length:   | 4680 |
| Score:   | 3360.50   | Matches:  | 609  |
| Percent Similarity:                                  | 89.42%    | Conservative:   | 50   |
| Best Local Similarity:                               | 82.63%    | Mismatches:   | 76   |
| Query Match:   | 84.24%    | Indels:   | 2    |
| DB:  | 1         | Gaps:   | 1    |
| US-09-807-802A-13 (1-736) x US-08-254-358-1 (1-4680) |           |   |      |
| Qy   | 1         | MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg    | 20   |
| Db   | 2203      | ATGGCTGCCGATGGTTATCTTCAGATTGGCTCGAGGACACTCTCTCTGAAGGAATAAGA     | 2262 |
| Qy   | 21        | GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp    | 40   |
| Db   | 2263      | CAGTGTGGAAAGCTCAAACCTGGCCACCAACCACCAAGCCGACAGCGGCAATAAGGAC      | 2322 |
| Qy   | 41        | AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp    | 60   |
| Db   | 2323      | GACAGCAGGGTCTTGCTTCTCGGTACAAAGTACCTCGGACCTTCAACGGGACTCGAC       | 2382 |
| Qy   | 61        | LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp       | 80   |
| Db   | 2383      | AAGGGAGCGGTCAACGAGGACGCGCGGCCCTCGAGCACGACAAAGCCTACGAC           | 2442 |
| Qy   | 81        | GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe    | 100  |
| Db   | 2443      | CGGCAGCTCGACAGCGGAGACAAACCGGTACCTCAAGTACAAACCAACCGCGGAGTTT      | 2502 |
| Qy   | 101       | GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln    | 120  |
| Db   | 2503      | CAGGAGCGCCTTAAAGAAGATACGTCTTTGGGGGCAACCTCGGACGAGCAGTCTTCCAG     | 2562 |
| Qy   | 121       | AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro       | 140  |
| Db   | 2563      | GCGAATAAGAGGGTCTTTGAACCTCTGGGCTGGTTGAGGAACCTGTTAAGACGGCTCCG     | 2622 |
| Qy   | 141       | GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly    | 160  |
| Db   | 2623      | GGAAAAAGAGGCGGTAGAGCACTCTCCTGTGGAGCCAGACTCTCTCGGGAACCGGA        | 2682 |
| Qy   | 161       | LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu    | 180  |
| Db   | 2683      | AAGCGGSCAGCAGCCTGCAAGAAAAAGATTGAATTTTGGTCAGACTTGAGACGCGAC       | 2742 |
| Qy   | 181       | SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro    | 200  |
| Db   | 2743      | TCAGTACCTGACCCCGCCTCTCGGACAGCCACCAAGCAGCCCTCTGGTCTGGGAAC        | 2802 |
| Qy   | 201       | ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly       | 220  |
| Db   | 2803      | AATACGATGGCTACAGGCAGTGGCGCAACCAATGGCAGACAATAACGAGGCGCGCGGA      | 2862 |
| Qy   | 221       | ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle    | 240  |
| Db   | 2863      | GTGGTAAATCTCCGGAATTTGGCAATTCGATTCACATGGATGGGCGACAGATCATC        | 2922 |
| Qy   | 241       | ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle    | 260  |
| Db   | 2923      | ACCACCAGCACCCGAACCTGGGCCCTGCCACCTACAAACCAACCACTCTACAAACAAAT     | 2982 |
| Qy   | 261       | SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp    | 280  |
| Db   | 2983      | TCCAGCCCAATCA---GGAGCCTCGAACGACAATCACTACTTTGGGTACAGCACCCCTTG    | 3039 |
| Qy   | 281       | GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu    | 300  |
| Db   | 3040      | GGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCAACCACTGACTGGCAAGACTC     | 3099 |
| Qy   | 301       | IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln    | 320  |
| Db   | 3100      | ATCAACAACAACACTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAACAATCAA | 3159 |

|    |      |   |      |
|----|------|---|------|
| Qy | 321  | VallysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThr  | 340  |
| Db | 3160 | GTCAAAGAGGTACGCAGAATGACGGTACGACGAGATTGCCAATAACCTTACCAGCAGC    | 3219 |
| Qy | 341  | ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln  | 360  |
| Db | 3220 | GTTCAGGTGTTTACTGACTCGGAGTACCAAGTCCCGTACGTCTCGGCTCGGCGCATCAA   | 3279 |
| Qy | 361  | GlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr     | 380  |
| Db | 3280 | GGATGCCTCCCGCGTTCAGCAGACGCTTCTTACGTCACAGTATGGATACCTCACC       | 3339 |
| Qy | 381  | LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro  | 400  |
| Db | 3340 | CTGAACAACGGGAGTCAGGCAGTAGGACGCTCTTCACTTTACTGCTGGAGTACTTCTCT   | 3399 |
| Qy | 401  | SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro  | 420  |
| Db | 3400 | TCTCAGATGCTGCTACCGGAACAACCTTTACTTTCAGCTACACTTTTGAGGACGTTCTCT  | 3459 |
| Qy | 421  | PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp  | 440  |
| Db | 3460 | TTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGAC  | 3519 |
| Qy | 441  | GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAsp  | 460  |
| Db | 3520 | CAGTACCTGTATTACTTGAGCAGAAACAACACTCCAAGTGGAAACCAACCGCAGTCAAGG  | 3579 |
| Qy | 461  | LeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuPro  | 480  |
| Db | 3580 | CTTCAGTCTTCTCAGGCGGAGCGAGTGACATTCGGGACCAAGTCTAGGAACTGGCTTCTCT | 3639 |
| Qy | 481  | GlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsn  | 500  |
| Db | 3640 | GGACCTGTATTACCGCCAGCAGCGAGTATCAAGACATCTCGGATAACAACAACAGTGAA   | 3699 |
| Qy | 501  | PheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnPro  | 520  |
| Db | 3700 | TACTCGTGAGCTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCG  | 3759 |
| Qy | 521  | GlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyVal  | 540  |
| Db | 3760 | GGGCGCGCCATGGCAAGCCACAAGGACGATGAAGAAAAGTTTTTCTCAGAGCGGGGTT    | 3819 |
| Qy | 541  | MetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIle  | 560  |
| Db | 3820 | CTCATCTTTGGGAAGCAAGGCTCAGAGAAAACAATGTGAACATTGAAAGGTCAATGATT   | 3879 |
| Qy | 561  | ThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrVal  | 580  |
| Db | 3880 | ACAGACGAAGAGGAATCGGAACAACCAATCCCGTGGCTACGGAGCAGTATGGTCTGTGA   | 3939 |
| Qy | 581  | AlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGly  | 600  |
| Db | 3940 | TCTACCAACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAACACACAAGGC  | 3999 |
| Qy | 601  | AlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAla  | 620  |
| Db | 4000 | GTTCTTCCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGGSCA  | 4059 |
| Qy | 621  | LysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeu  | 640  |
| Db | 4060 | AAGATTCCACACACGGACGGACATTTTCAACCCCTCTCCCTCATGGGTGGATTCTGGACTT | 4119 |
| Qy | 641  | LysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAla     | 660  |
| Db | 4120 | AAACACCTCTCTCCACAGATTCTCATCAAGAACACCCCGTACCTCGGAATCTCTTCGACC  | 4179 |
| Qy | 661  | GluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGln-ValSe  | 680  |
| Db | 4180 | ACCTTCAGTGGGCAAGTTTGCTTCTTTCATCACACAGTACTCCACGGGACACGGTCAG    | 4239 |
| Qy | 680  | rValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGl  | 700  |



```

Db      4240 CGTGGAGATCGAGTGGGAGCTGCAGAAGGAAACAGCAACCGTGAATCCCGAAATTC A 4299
QY      700 nTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLe 720
Db      4300 GTACACTTCCAACACTACAACAAGTCTGTTAATCGTGGACTTACCGTGGATACATAATGGCGT 4359
QY      720 uTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736
Db      4360 GTATTACAGAGCCTCGCCCATTTGGCACCAGATACCTTGACTCGTAATCTG 4408

RESULT 6
US-08-475-391-1
; Sequence 1, Application US/08475391
; Patent No. 5786211
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; METHOD OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshhall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,391
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/254,358
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 578621land, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31975
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-475-391-1

Alignment Scores:
Pred. No.:          9,71e-315          Length:          4680
Score:             3360.50             Matches:          609
Percent Similarity: 89.42%             Conservative:      50
Best Local Similarity: 82.63%           Mismatches:       76
Query Match:       84.24%               Indels:           2
DB:                1                   Gaps:            1

US-09-807-802A-13 (1-736) x US-08-475-391-1 (1-4680)
QY      1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20
Db      2203 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACACTCTCTCTGAAGGAATAAGA 2262
QY      21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp 40
Db      2263 CAGTGGTGAAGCTCAAAACCTGGCCCAACCAACCAAGCCCGCAGAGCGGCATAGGAC 2322
```

```

QY      41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyPropheAsnGlyLeuAsp 60
Db      2323 GACAGACAGGGGTCTGTGCTTCTCTGGGTACAGTACCTCGGACCCCTTCAACGGACTCGAC 2382
QY      61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80
Db      2383 AAGGGAGAGCCGGTCAACGAGGACAGCGCCGCCCTCGAGCAGCAAAAGCCTACGAC 2442
QY      81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100
Db      2443 CGGCAGCTCGACAGCGGAGACAACCCCGTACCTCAAGTACAACCAACCGCCGACGCGAGTTT 2502
QY      101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120
Db      2503 CAGGAGCGCCTTAAAGAAGATACGTCTTTTGGGGCAACCTCGGACGAGCAGTCTTCCAG 2562
QY      121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro 140
Db      2563 GCGAAAAAGAGGGTCTTGAACCTCTGGGCCCTGGTTGAGGAACCTGTTAAGACGGCTCCG 2622
QY      141 GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly 160
Db      2623 GGAATAAAGAGGCGGTAGAGCACTCTCTGTGAGCCAGACTCTCTCTCGGGAACCGGA 2682
QY      161 LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180
Db      2683 AAGCGGGCCAGCAGCCTGCAAGAAAAAGATTGAATTTTGGTCAGACTGGAGACGACGAC 2742
QY      181 SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro 200
Db      2743 TCAGTACCTGACCCCGAGCCTCTCGGACAGCCACCGAGCCCTCTCTGCTGGGAACT 2802
QY      201 ThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly 220
Db      2803 AATACGATGGCTACAGGCAGTGGCGCACCAATGCGCAGACAATAACGAGGGCGCCGACGGA 2862
QY      221 ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle 240
Db      2863 GTGGGTAAATTCCTCCGAAATTTGGCATTTGCGATTCCGATGATGGGCGACAGAGTCATC 2922
QY      241 ThrThrSerThrArgThrTrpAlaLeuProThrTrpAsnAsnHisLeuTyrLysGlnIle 260
Db      2923 ACCACGACACCCGAACTGGGCCCTGCCACCTACAAACAACCACTCTACAAACAAT 2982
QY      261 SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp 280
Db      2983 TCCAGCCAATCA ---GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTGG 3039
QY      281 GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu 300
Db      3040 GGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCCACCACGTGACTGGCAAAGACTC 3099
QY      301 IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln 320
Db      3100 ATCAACAACAACGTGGGATTCGACCCCAAGAGACTCAACTTCAAGCTCTTTTAACATTCAA 3159
QY      321 ValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThr 340
Db      3160 GTCAAAAGAGGTCCACGCAGAAATGACGGTACGACGACGATGGCCAATAACCTTACCAGCAG 3219
QY      341 ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln 360
Db      3220 GTTCAGGTGTTTACTGACTCGGAGTACCAGCTCCCGTACGCTCTCGGCTCGGCGCATCAA 3279
QY      361 GlyCysLeuProPropheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr 380
Db      3280 GGATGCTCCCGCCGTTCCACGACAGACGTCTTTCATGGTGCCACAGTATGGATACCTCACC 3339
QY      381 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 400
Db      3340 CTGAACAACGGGAGTCAGGCAGTAGGACGCTCTTTCATTTTACTGCCTGGAGTACTTTTCT 3399
```

```
QY 401 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluValPro 420
Db 3400 TCTCAGATGCTGCGTACCGAAACAACCTTTACCTTCAGCTACACTTTTGAGCAGTTCCT 3459
QY 421 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 440
Db 3460 TTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGAC 3519
QY 441 GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAsp 460
Db 3520 CAGTACCTGTATTACTTGAGCAGAGAACAACACTCCAAGTGAACCAACCACCGCAGTCAAGG 3579
QY 461 LeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuPro 480
Db 3580 CTTTCAGTTTCTCAGSCCGGAGCGAGTGACATTCGGGACCAGTCTAGGAACCTGCTCCT 3639
QY 481 GlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsn 500
Db 3640 GGACCCTGTTACCGCCAGCAGCGAGTATCAAAAGACATCTGCGGATAACAACACAGTGAA 3699
QY 501 PheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnPro 520
Db 3700 TACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCG 3759
QY 521 GlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyVal 540
Db 3760 GGGCCCGCCATGGCAAGCCACAGGACGATGAAGAAAGTTTTTCCTCAGAGCGGGTT 3819
QY 541 MetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIle 560
Db 3820 CTCATCTTTGGGAAGCAAGGCTCAGAGAAAACAAATGTGAACATTGAAAAGTTCATGATT 3879
QY 561 ThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrVal 580
Db 3880 ACAGACGAAGAGGAAATCGGAACAACCAATCCCGTGGCTACGGAGCAGTATGGTTCGTGA 3939
QY 581 AlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGly 600
Db 3940 TCTACCAACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAACACACAAGGC 3999
QY 601 AlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAla 620
Db 4000 GTTCTTCCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGCA 4059
QY 621 LysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeu 640
Db 4060 AAGATTCCACACACGAGCAGACATTTTCAACCCCTCTCCCTCATGGGTGGATTTCGACTT 4119
QY 641 LysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAla 660
Db 4120 AAACACCTCTCTCCACAGATTCTCATCAAGAAACACCCCGGTACCTGCGAATCCTTCGACC 4179
QY 661 GluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGln-ValSe 680
Db 4180 ACCTTCAGTGGCGCAAGTTTGCTTCTTCATCACACAGTACTCCACGGGACACGGTTCAG 4239
QY 680 rValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValG1 700
Db 4240 CGTGGAGATCGAGTGGAGCTGCAGAAGGAAACAGCAACGCTGGAATCCCGAAATTCA 4299
QY 700 nTyrThrSerAsnTyrAlaLysSerAlaAsnValAaspPheThrValAspAsnAsnGlyLe 720
Db 4300 GTACACTTCCAACTACAACAAGTCTGTTAATCGTGGACTTACCGTGATACTAATGGCGT 4359
QY 720 uTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736
Db 4360 GTATTACAGAGCCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTG 4408
```

RESULT 7

US-08-709-609-1

; Sequence 1, Application US/08709609

; Patent No. 5858775

; GENERAL INFORMATION:

```
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,609
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: NO. 5858775and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31975
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-709-609-1
```

Alignment Scores:  
Pred. No.: 9,71e-315 Length: 4680  
Score: 3360.50 Matches: 609  
Percent Similarity: 89.42% Conservative: 50  
Best Local Similarity: 82.63% Mismatches: 76  
Query Match: 84.24% Indels: 2  
DB: 2 Gaps: 1

US-09-807-802A-13 (1-736) x US-08-709-609-1 (1-4680)

```
QY 1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20
Db 2203 ATGGCTGCCGATGGTTATCTTCCAGATGGCTCGAGGACACTCTCTCTGAAGGAATAAGA 2262
QY 21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnLysGlnAsp 40
Db 2263 CAGTGGTGGAAAGCTCAAACCTGGCCCAACACCACCAAGCCCGCAGAGCGGCATAAGGAC 2322
QY 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyPropheAsnGlyLeuAsp 60
Db 2323 GACAGCAGGGGTCTTGTTCTTCTGGGTACAAGTACCTCGGACCCCTTCAACGGACTCGAC 2382
QY 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80
Db 2383 AAGGGAGAGCCGGTCAACGAGGCGAGCAGCGCGCGGCCCTCGAGCAGCAAAAGCCTACGAC 2442
QY 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100
Db 2443 CGGCAGCTCGACAGCGGAGACAACCCGCTACCTCAAGTACAACCCGCGCGCGGAGTTT 2502
QY 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120
Db 2503 CAGGAGCGCCTTAAGAAGATACGTCTTTTGGGGGCAACCTCGGACGAGCAGTCTTCCAG 2562
QY 121 AlaLysLysArgValLeuGluProLeuGluValGluGlyAlaLysThrAlaPro 140
```

Db 2563 GCGAAAAAGAGGGTTCTTGAACCTCTGGGCCCTGGTTGAGGAACCTGTTAAGACGGCTCCG 2622

Qy 141 GlyLysArgProValGluGlnSerProGlnGluProAspSerSerSerGlyLeGly 160  
|||||  
Db 2623 GGAAGAAAGAGCGCGGTAGAGCACTCTCTGTGGAGCCAGACTCTCTCGGGAACCGGA 2682

Qy 161 LysThrGlyGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180  
|||||  
Db 2683 AAGCGGGCCAGCAGCCTCGAAGAAAAAGATTGAATTTTGGTCAGACTGGAGACGCAGAC 2742

Qy 181 SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro 200  
|||||  
Db 2743 TCAGTACCTGACCCCGCCTCTCGACAGCCACCCAGCCCTCTGGTCTGGGAAC 2802

Qy 201 ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly 220  
|||||  
Db 2803 AATACGATGGCTACAGGCAGTGGCGCACCAATGGCAGACATAACGAGGGCGCGACGGA 2862

Qy 221 ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle 240  
|||||  
Db 2863 GTGGGTAAATCTCCGGAATTTGGCATTTGGATTCCACATGGATGGCGCAGAGTCATC 2922

Qy 241 ThrThrSerThrArgThrTrpAlaLeuProThrTrpAsnAsnHisLeuThrLysGlnIle 260  
|||||  
Db 2923 ACCACGACCCCGAACCTGGGCCCTGCCCCACCTACAACAACACACCTCTACAAACAAT 2982

Qy 261 SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp 280  
|||||  
Db 2983 TCCAGCCCAATCA---GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTGG 3039

Qy 281 GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu 300  
|||||  
Db 3040 GGGTATTTTGAATTCACAGATTCCACTGCCACTTTTCCACCAGTGAATGGCAAGACTC 3099

Qy 301 IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln 320  
|||||  
Db 3100 ATCAACAACAACACTGGGATTCGACCCCAAGAGACTCAACTTCAAGCTCTTTAACAATTCAA 3159

Qy 321 ValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThr 340  
|||||  
Db 3160 GTCAAAGAGGTACGCAGAAATGACGGTACGACGAGATGGCCAATAACCTTACCAGCACG 3219

Qy 341 ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln 360  
|||||  
Db 3220 GTTCAGGTGTTTACTGACTCGGAGTACCACTCCGTCAGTCCCTCGGCTCGGCGCATCAA 3279

Qy 361 GlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr 380  
|||||  
Db 3280 GGATGCCTCCCGCGGTTCCAGCAGACGCTCTCATGGTGGCCACAGTATGGATACCTCACC 3339

Qy 381 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 400  
|||||  
Db 3340 CTGAACAACGGGAGTCAGGCAGTAGGACGCTCTTCACTTTACTGCTGGAGTACTTTTCT 3399

Qy 401 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro 420  
|||||  
Db 3400 TCTCAGATGCTGCTACCGGAACAACCTTTACTTACCTCAGCTACACTTTTGAGGACGTTCT 3459

Qy 421 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 440  
|||||  
Db 3460 TTCCACAGCAGCTACGCTCACAGCCAGAGTCTGACCGCTCTCATGATCTCTCATCGAC 3519

Qy 441 GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAsp 460  
|||||  
Db 3520 CAGTACCTGTATTACTTGAGCAGAAACAACACTCCAAGTGGAAACCAACCCAGCTCAAGG 3579

Qy 461 LeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuPro 480  
|||||  
Db 3580 CTTTCAGTTTTCTCAGCGCGGAGCGAGTGACATTTCGGGACCACTAGGAATGGCTTCTCT 3639

Qy 481 GlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnSerAsn 500  
|||||  
Db 3640 GGACCCCTGTTACCGCCAGCAGCGAGTATCAAGACATCTGCGGATACAACAACAGTGA 3699

Qy 501 PheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnPro 520  
|||||  
Db 3700 TACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCG 3759

Qy 521 GlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyVal 540  
|||||  
Db 3760 GGGCCCGCATGGCAAGCCACAAGCAGATGAAGAAAAAGTTTTTCTCCTCAGAGCGGGGTT 3819

Qy 541 MetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIle 560  
|||||  
Db 3820 CTCATCTTTGGGAAGCAAGGCTCAGAGAAAAACAATGTGAACATTGAAAAAGGTCTGATT 3879

Qy 561 ThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrVal 580  
|||||  
Db 3880 ACAGACGAAGAGGAATCGGAACAACCAATCCCGTGGCTACGGAGCAGTATGTTCTGTA 3939

Qy 581 AlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGly 600  
|||||  
Db 3940 TCTACCAACCTCCAGAGAGGCAACAGACAAGCAGTACCGCAGATGTCAACACACAAGGC 3999

Qy 601 AlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAla 620  
|||||  
Db 4000 GTTCTTCCAGGCATGTTCTGGCAGGACAGAGATGTGTACTTTCAGGGGCCCATCTGGCA 4059

Qy 621 LysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeu 640  
|||||  
Db 4060 AAGATTCCACACACGAGCGGACATTTTCAACCCCTCTCCCTCATGGTGGATTCCGACTT 4119

Qy 641 LysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAla 660  
|||||  
Db 4120 AAACACCTCTCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAATCCTTCGACC 4179

Qy 661 GluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGln-ValSe 680  
|||||  
Db 4180 ACCTTCAGTGGGCAAGTTTGCTTCTTTCATCATCACAGTACTCCACGGGACACGGTCAG 4239

Qy 680 rValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValG 700  
|||||  
Db 4240 CGTGAGATCGAGTGGAGCTGCAGAGGAAACAGCAACGCTGGAATCCCGAAATTC 4299

Qy 700 nTyrThrSerAsnTyrAlaLysSerAlaAsnValaPheThrValAspAsnAsnGlyLe 720  
|||||  
Db 4300 GTACACTTCCAACACTACAAGTCTGTTAATCGTGGACTTACCGTGGATATACTAATGGCGT 4359

Qy 720 uTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736  
|||||  
Db 4360 GTATTTCAGAGCCTCGCCCATTCGACCAGATACCTGACTGCTGTAATCTG 4408

RESULT 8  
PCT-US95-07178-1  
; Sequence 1, Application PC/TUS9507178  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Philip R.  
; TITLE OF INVENTION: Adeno-Associated Virus Materials and  
; TITLE OF INVENTION: Methods  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 S. Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/07178  
; FILING DATE:  
; CLASSIFICATION:





```
QY 601 AlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAla 620
Db 4000 GTTCTTCCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGGCA 4059
QY 621 LysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeu 640
Db 4060 AAGATTCCACACACGCGGACATTTTCACCCCTCTCCCTCATGGTGGATTCCGACTT 4119
QY 641 LysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAla 660
Db 4120 AAACACCTCTCTCCACAGATTCTCATCAAGAAACACCCCGGTACCTGCGAATCCTTCGACC 4179
QY 661 GluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGln-Valse 680
Db 4180 ACCTTCAGTGGCGCAAGTTGCTTCTTCATCAACAGTACTCCACGGGACACGGTCAG 4239
QY 680 rValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValG1 700
Db 4240 CGTGGAGATCGAGTGGGAGCTGCAGAGGAAACAGCAACCGTGGAAATCCCGAAATTC 4299
QY 700 nTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLe 720
Db 4300 GTACACTTCCAACTACAAAGTCTGTTAATCGTGGACTTACCGTGGATACCTAAATGCGT 4359
QY 720 uTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736
Db 4360 GTATTGAGAGCTCGCCCAATGGCACCAGATACCTGACTCGTAATCTG 4408
```

RESULT 9

```
US-09-438-268-2
; Sequence 2, Application US/09438268
; Patent No. 6491907
; GENERAL INFORMATION:
; APPLICANT: Rabinowitz, Joseph E.
; APPLICANT: Samulski, Richard J
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 5470-186
; CURRENT APPLICATION NUMBER: US/09/438,268
; CURRENT FILING DATE: 1999-11-10
; EARLIER APPLICATION NUMBER: 60/107,840
; EARLIER FILING DATE: 1998-11-10
; EARLIER APPLICATION NUMBER: 60/123,651
; EARLIER FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 8151
; TYPE: DNA
; ORGANISM: Virus
US-09-438-268-2
```

```
Alignment Scores:
Pred. No.: 3.67e-283 Length: 8151
Score: 3038.50 Matches: 552
Percent Similarity: 83.85% Conservative: 66
Best Local Similarity: 74.90% Mismatches: 108
Query Match: 76.17% Indels: 11
DB: 4 Gaps: 3
```

US-09-807-802A-13 (1-736) x US-09-438-268-2 (1-8151)

```
QY 1 MetaAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20
Db 2133 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACACTCTCTCTGAAGGATAAGA 2192
QY 21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp 40
Db 2193 CAGTGGTGAAGCTCAAACCTGGCCACCACCACCAAGGCCGAGCGGCGCATAGGAC 2252
QY 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyPropheAsnGlyLeuAsp 60
```

```
Db 2253 GACAGCAGGGGTCTTGTGCTTCTCGGTACAAGTACCTCGGACCTTCAACGGACTCGAC 2312
QY 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80
Db 2313 AAGGGAGAGCCGGTCAACGAGGACAGCGCCGCGGCCCTCGAGCACGACAAGGCCTACGAC 2372
QY 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100
Db 2373 CAGCAGCTCAAGGCCGGTGACAAACCTTACCTCAAGTACAACACGCGGACGCGGAGTTC 2432
QY 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120
Db 2433 CAGCAGCGGCTTCAGGGCGACACATCGTTTGGGGCAACCTCGGCAGAGCAGTCTTCCAG 2492
QY 121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro 140
Db 2493 GCCAAAAGAGGGTCTTGAACCTCTTGGTCTGTTGAGCAAGCGGGTGAGACGGCTCCT 2552
QY 141 GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly 160
Db 2553 GGAAGAGAGACCGTTGATTGAATCCCCCAGCAGCCCGACTCTCTCCACGGGTATCGGC 2612
QY 161 LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPhe---GlyGlnThrGlyAspSer 179
Db 2613 AAAAAAGCAAGCAGCGCGCTAAAAAAGAGCTCGTTTTCGAAGACGAAACTGGAGCAGGC 2672
QY 180 GluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGly 199
Db 2673 GACGGACCCCTGAGGGATCAACTTCCGGA-----GCCATGTCTGTAT 2714
QY 200 ProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAsp 219
Db 2715 GACAGTGAGATCGGTGCAGCAGCTGGCGAGCTGCAGTCGAGGCGGCAAGGTGCCGAT 2774
QY 220 GlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgVal 239
Db 2775 GGAGTGGTAAATGCCTCGGTGATTGGCATTCGATTCCACCTGGTCTGAGGGCCACGTC 2834
QY 240 IleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGln 259
Db 2835 ACGACCACCAAGCAGCAGAACCTGGGTCTTGCCCCACCTACAACAACCACTCTACAAGCGA 2894
QY 260 IleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrPro 279
Db 2895 CTCGGAGAG-----AGCCTGCAGTCCAACACCTACAACGGATTCTCCACCCCC 2942
QY 280 TrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArg 299
Db 2943 TGGGGATACTTTGACTTCAACCGCTTCCACTGCCACTTCTTCAACCACTGCTGAGTGGCAGCGA 3002
QY 300 LeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIle 319
Db 3003 CTCATCAACAACAACCTGGGCGATGCGACCCCAAGCCATGCGGGTCAAAATCTTCAACATC 3062
QY 320 GlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSer 339
Db 3063 CAGGTCAAGGAGGTACAGACGTGCAACGCGGAGAGACAACGGTGGCTAATAACCTTACCAGC 3122
QY 340 ThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHis 359
Db 3123 ACGTTTCAGATCTTTGGCGACTCGTCGTACGAACCTGCCGTACGTCTCGGCTCGGCGCAT 3182
QY 360 GlnGlyCysLeuProPropheProAlaAspValPheMetIleProGlnTyrGlyTyrLeu 379
Db 3183 CAAGGATGCCCTCCCGCGTTCAGCAGACGCTTTCATGGTGCCACAGTATGGATACCTC 3242
QY 380 ThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhe 399
Db 3243 ACCCTGAACAACGGGAGTCAGGCAGTAGGACGCTTCTTCTTACTGCTGAGTACTTT 3302
QY 400 ProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluVal 419
```

```
Db 3303 CCTTCTCAGATGCTGCTACCGGAAACAACTTTACCTTCAGCTACACTTTTGAGGACGTT 3362
QY 420 PropHeHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIle 439
Db 3363 CCTTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATC 3422
QY 440 AspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLys 459
Db 3423 GACCAGTACCTGTATTACTTGAGCAGAAACAACACTCCAAGTGAACCAACCACGAGTCA 3482
QY 460 AspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeu 479
Db 3483 AGCTTCAGTTTTCTCAGGCGGAGCGAGTGACATTCGGGACCAGTCTAGGAATGGCTT 3542
QY 480 ProGlyProCysTyrArgGlnArgValSerLysThrLysThrAspAsnAsnAsnSer 499
Db 3543 CCTGGACCTGTTACCGCCAGCAGCGAGTATCAAAGACATCTGCGGATAACAACAACAGT 3602
QY 500 AsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleAsn 519
Db 3603 GAATACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAAT 3662
QY 520 ProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGly 539
Db 3663 CCGGCGCCGCCATGGCAAGCACAAGGACGATGAAGAAAAAGTTTTTTCCTCAGAGCGGG 3722
QY 540 ValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMet 559
Db 3723 GTTCTCATCTTTGGGAAGCAAGGCTCAGAGAAACAATGTGAACATTGAAAGGTCTATG 3782
QY 560 IleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThr 579
Db 3783 ATTACAGACGAAGAGGAATCGGAACAACCAATCCCCGTGGTACGGAGCAGTATGTTCT 3842
QY 580 ValAlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMet 599
Db 3843 GTATCTACCAACCTCCAGAGGCAACAGACAAGCAGTACCAGCATGTCAACACACAA 3902
QY 600 GlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrp 619
Db 3903 GGCCTTCTCCAGGCATGGTCTGGCAGGACAGAGATGTACCTTCAGGGGCCATCTGG 3962
QY 620 AlalysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGly 639
Db 3963 GCAAAGATTCCACACACGACGACATTTTCAACCCCTCTCCCTCATGGTGGATTCCGA 4022
QY 640 LeuLysAsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProPro 659
Db 4023 CTTAAACACCCCTCCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAATCCTCG 4082
QY 660 AlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnVal 679
Db 4083 ACCACCTTCAGTCGGGCAAGTTGCTTCTTCATCACACAGTACTCCACGGGACAGGTC 4142
QY 680 SerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluVal 699
Db 4143 AGCGTGAGATCGAGTGGAGCTGCAGAAGGAAAAACAGCAAAACGCTGGAATCCCGAAATT 4202
QY 700 GlnTyrThrSerAsnTyrAlalysSerAlaAsnValAspPheThrValAspAsnAsnGly 719
Db 4203 CAGTACACTTCCAACCTACAACAAGTCTGTTAATCGTGGACTTACCGTGATATAATGGC 4262
QY 720 LeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736
Db 4263 GTGTATTACAGAGCCTCGCCCCATTGGCACCAGATACCTGACTCGTAATCTG 4313
```

RESULT 10

```
US-09-532-594B-5
; Sequence 5, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
```

```
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2208
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: synthetic construct
; NAME/KEY: misc_feature
; LOCATION: 750
; OTHER INFORMATION: n = g, a, c or t(u)
; NAME/KEY: misc_feature
; OTHER INFORMATION: AAV4 caspid protein VP1 gene
; US-09-532-594B-5
```

```
Alignment Scores:
Pred. No.: 1.17e-230 Length: 2208
Score: 2486.50 Matches: 475
Percent Similarity: 74.33% Conservative: 81
Best Local Similarity: 63.50% Mismatches: 161
Query Match: 62.33% Indels: 31
DB: 4 Gaps: 10
```

US-09-807-802A-13 (1-736) x US-09-532-594B-5 (1-2208)

```
QY 4 AspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArgGluTrpTrp 23
Db 7 GACGGTTACCTTCAGATTGGCTAGAGGACAACCTCTCTGAAGGCGTTCGAGAGTGGTGG 66
QY 24 AspLeuLysProGlyAlaProLysProLysProLysAlaAsnGlnGlnLysGlnAspAspGlyArg 43
Db 67 GCGCTGCAACCTGGAGCCCCCTAAACCCCAAGGCAAAATCAACATCAGGACAACGCTCGG 126
QY 44 GlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAspLysGlyGlu 63
Db 127 GGTCTTGTGCTTCGGGTTTACAAATACCTCGGACCCGCGCAACGACTCGACAGGGGAA 186
QY 64 ProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAspGlnGlnLeu 83
Db 187 CCCGTCAACGACGCGACGCGGCGAGCCCTCGAGCACGACAAGGCCCTACGACCGAGCTC 246
QY 84 LysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPheGlnGluArg 103
Db 247 AAGCCCGGTGACACCCCTACCTCAAGTACACCAACCGCGGAGTTCGAGCAGCGG 306
QY 104 LeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGlnAlaLysLys 123
Db 307 CTTCAGGGCGACACATCGTTTGGGGGCAACCTCGGCAGAGCAGTCTTCCAGGCCAANAAG 366
QY 124 ArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaProGlyLysLys 143
Db 367 AGGGTTCTTGAACTCTTGGTCTGGTTGAGCAAGCGGGTGAGACGGCTCTCTGGAAAGAAG 426
QY 144 ArgProValGlnSerProGlnGluProAspSerSerSerGlyIleGlyLysThrGly 163
Db 427 AGACCGTTGATTGAATCCCCCAGCAGCCCGACTCCTCCACGGGTATCGGCAAAAAAGGC 486
QY 164 GlnGlnProAlaLysLysArgLeuAsnPhe---GlyGlnThrGlyAspSerGluSerVal 182
Db 487 AAGCAGCCCGCTAAAAAAGAAGCTCGTTTTTCGAAGACGAACACTGGAGCAGCGCGGACCCC 546
QY 183 ProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyProThrThr 202
Db 547 CCTGAGGGATCAACTTCCGGA-----GCCATGTCTGTGATGACAGTGAG 588
QY 203 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 222
```



Db 589 ATGCGTGCGAGCAGCTGGCGGAGCTGCAGTCGAGGSGGACAAAGTGCCTGAGTGGGT 648  
Qy 223 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 242  
Db 649 AATGCTCGGGTGATGGCATTTGGCATTCACCTGGTCTGAGGGCCACGTCACGACCAC 708  
Qy 243 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 262  
Db 709 AGCACCAGAACCTGGGTCTTGCCCACTTACCAACCAACCACTTNTACAAGCGACTCGGAGAG 768  
Qy 263 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 282  
Db 769 -----AGCTGCAGTCCCAACACCTACAACCGATTCTCCACCCCTGGGGATAC 816  
Qy 283 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 302  
Db 817 TTTGACTTCAACCGCTTCCACTGCCACTTCTCACCACGTCGACGCGACTCATCAAC 876  
Qy 303 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 322  
Db 877 AACAACTGGGCGATCGACCCCAAGCCATGCGGGTCAAAATCTTCAACATCCAGTCAAG 936  
Qy 323 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 342  
Db 937 GAGGTACGACGTCGAACGGCGAGACACGGTGGTAAATAACCTTACCAGCACGGTTCAG 996  
Qy 343 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 362  
Db 997 ATCTTTGCGGACTCGTCGTACGAACCTGCCGTACGTGATGATGCGGGTCAAGAGGGCAGC 1056  
Qy 363 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyr-----Leu 379  
Db 1057 CTGCTCTCTTTCCCAACGACGCTCTTTATGTGTGCCCCAGTACGGCTACTGTGGACTGGT 1116  
Qy 380 ThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhe 399  
Db 1117 ACCGGCAACACTTCGCAGCAACAGACTGCAGCAAAATGCCTTCTACTTCGCTGGAGTACTTT 1176  
Qy 400 ProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluVal 419  
Db 1177 CCTTCGAGATGTCGGGACTGGCAACAACTTTGAAATTACGTACAGTTTGAGAAGGTG 1236  
Qy 420 ProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIle 439  
Db 1237 CCTTCCACTCGATGTACGCGCACAGCCAGAGCTGGACCGGCTGATGAACCCCTCTCATC 1296  
Qy 440 AspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsn-----GlnSerGlySer 455  
Db 1297 GACCAGTACCTGTGGGACTGCAATCGACCACCAACCGGAAACCCCTGAAATCGCGGACT 1356  
Qy 456 AlaGlnAsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnPro 475  
Db 1357 GCCACCACCAAC-----TTTACCAAGCTGCGGCTTACCAACTTTTCCAACTTTAAA 1407  
Qy 476 LysAsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAsp 495  
Db 1408 AAGAACTGGCTGCCCGGCTTCAATCAAGCAGCAGGGCTTCTCAAAAGACTGCC----- 1461  
Qy 496 AsnAsnAsnSerAsnPheThrTrpThrGlyAlaSer-----LysTyr----- 509  
Db 1462 AATCAAAACTACAAGATCCCTGCCACCGGGTTCAGACAGTCTCATCAATACGAGACGCAC 1521  
Qy 510 ---AsnLeuAsnGlyArgGluSerIleIleAsnProGlyThrAlaMetAlaSerHisLys 528  
Db 1522 AGCACTCTGGACGGAAGATGGAGTGCCTGACCCCGGACCTCCAATGGCCACGGCTGGA 1581  
Qy 529 AspAspGluAspLysPhePheProMetSerGlyValMetIlePheGlyLysGluSerAla 548  
Db 1582 CCTGCGGACAGCAAG---TTCAGCAACAGCCAGCTCATCTTTGCGGGGCTAAACAGAAC 1638  
Qy 549 GlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAspGluGluIleLysAla 568

Db 1639 GGCAACACGGCCACCGTACCCGGGACTCTGTATCTTACCTCTGAGGAGGAGCTGGCAGCC 1698  
Qy 569 ThrAsnProValAlaThrGluArgPheGlyThrValAlaValAsnPheGlnSerSerSer 588  
Db 1699 ACCAACGCCACCGATACGACATGTGGGGCAACCTTACCTGGCGGTGACCAAGCAACAGC 1758  
Qy 589 ThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeuProGlyMetValTrpGln 608  
Db 1759 AACCTGCCGACCGTGGACAGACTGACAGCCCTTGGGAGCCGTGCCTGGAATGGTCTGCAA 1818  
Qy 609 AspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIleProHisThrAspGlyHis 628  
Db 1819 AACAGAGACATTTACTACAGGGTCCCATTTGGGCCAAGATTCTCATACCGATGGACAC 1878  
Qy 629 PheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsnProProGlnIleLeu 648  
Db 1879 TTTCAACCCCTCACCGCTGATTTGGTGGTGGTGGCTGAAACACCCGCTCTCTCAAAATTTT 1938  
Qy 649 IleLysAsnThrProValProAlaAsnProProAlaGluPheSerAlaThrLysPheAla 668  
Db 1939 ATCAAGAAACACCCCGTACCTGCGAATCCTGCAACGACCTTCACTCTCCGCTAAAC 1998  
Qy 669 SerPheIleThrGlnTyrSerThrGlyGlnValSerValGluIleGluTrpGluLeuGln 688  
Db 1999 TCCTTCATTACTCAGTACAGCACTGGCCAGGTGTCGGTGCAGATTGACTGGGAGATCCAG 2058  
Qy 689 LysGluAsnSerLysArgTrpAsnProGluValGlnTyrThrSerAsnTyrAlaLysSer 708  
Db 2059 AAGGAGCGGTCCAAACCGCTGGAACCCCGAGGTCCAGTTTACCTCCAACACGACAGCAA 2118  
Qy 709 AlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThrGluProArgProIleGly 728  
Db 2119 AACTCTCTGTGGGCTCCGATCGGGCTGGGAAATACACTGAGCCTAGGGCTATCGGT 2178  
Qy 729 ThrArgTyrLeuThrArgProLeu 736  
Db 2179 ACCCGCTACCTCACCCACCACCTG 2202

RESULT 11

US-09-438-268-1/c  
; Sequence 1, Application US/09438268  
; Patent No. 6491907  
; GENERAL INFORMATION:  
; APPLICANT: Rabinowitz, Joseph E.  
; APPLICANT: Samuleki, Richard J  
; APPLICANT: Xiao, Weidong  
; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING  
; FILE REFERENCE: 5470-186  
; CURRENT APPLICATION NUMBER: US/09/438,268  
; CURRENT FILING DATE: 1999-11-10  
; EARLIER APPLICATION NUMBER: 60/107,840  
; EARLIER FILING DATE: 1998-11-10  
; EARLIER APPLICATION NUMBER: 60/123,651  
; EARLIER FILING DATE: 1999-03-10  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 7214  
; TYPE: DNA  
; ORGANISM: Virus  
US-09-438-268-1

Alignment Scores:  
Pred. No.: 8.34e-230 Length: 7214  
Score: 2486.50 Matches: 475  
Percent Similarity: 74.33% Conservative: 81  
Best Local Similarity: 63.50% Mismatches: 161  
Query Match: 62.33% Indels: 31  
DB: 4 Gaps: 10

US-09-807-802A-13 (1-736) x US-09-438-268-1 (1-7214)

QY 4 AspGlyTyrLeuProAspTyrLeuGluAspAsnLeuSerGluGlyIleArgGluTyrTrp 23  
DB 2977 GACGGTTACCTTCAGATTGGCTAGAGGACAACTCTCTGAAGCGGTTTCAGAGTGGTG 2918  
QY 24 AspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAspAspGlyArg 43  
DB 2917 GCGCTGCAACCTGGAGCCCTAAACCCAAAGCAATCAACAATCAGGACAAACGCTCGG 2858  
QY 44 GlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAspLysGlyGlu 63  
DB 2857 GGTCTTGCTTCGGGTTACAAATACCTCGACCCGGCAACGGACTCGACAAGGGGAA 2798  
QY 64 ProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAspGlnGlnLeu 83  
DB 2797 CCCGTCAACGCGGACGCGGACGCCCTCGAGACGACAAGGCCCTACGACCGAGCTC 2738  
QY 84 LysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPheGlnGluArg 103  
DB 2737 AAGGCCGGTGACAAACCCCTACCTCAAGTACAAACACCGCCGACGGGAGTTCAGACGCGG 2678  
QY 104 LeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGlnAlaLysLys 123  
DB 2677 CTTCAGGGCGACACATCGTTGGGGCAACCTCGGCAGAGCAGTCTTCCAGGCCAAAGAAG 2618  
QY 124 ArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaProGlyLysLys 143  
DB 2617 AGGTTCTTGAACCTCTTGGTCTGTTGAGCAAGCGGTGAGACGGCTCTTGAAAGAAG 2558  
QY 144 ArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGlyLysThrGly 163  
DB 2557 AGACCGTTGATTGAATCCCCCAGCAGCCCCGACTCTCCACGGGTATCGGCAAAAAGGC 2498  
QY 164 GlnGlnProAlaLysLysArgLeuAsnPhe---GlyGlnThrGlyAspSerGluSerVal 182  
DB 2497 AAGCAGCCGGCTAAAGAGAGCTCGTTTCGAGACGAAACTGGAGCAGCGGACGCC 2438  
QY 183 ProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyProThrThr 202  
DB 2437 CCTGAGGGATCAACTTCCGA-----GCCATGCTGTGATGACAGTGAG 2396  
QY 203 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 222  
DB 2395 ATGCGTGCAGCAGCTGGCGAGCTGCAGTCGAGGGCGGACAAGGTGCCGATGGAGTGGGT 2336  
QY 223 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 242  
DB 2335 AATGCTCGGGTGATTGGGATTGCGATTCCACCTGTGTGAGGGCCACGTGACGACCACC 2276  
QY 243 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 262  
DB 2275 AGCACCAGAACCTGGGTCTTGCCACCTACAACAACCACTCTACAAGCGACTCGGAGAG 2216  
QY 263 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 282  
DB 2215 -----AGCCTGAGTCCAACACCTACAACGGAATCTCCACCCCTGGGGATAC 2168  
QY 283 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 302  
DB 2167 TTGACTTCAACCGCTTCCACTGCCACTTCTCACCAACGTAATCTTCAACATCCAGGTCAAC 2108  
QY 303 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 322  
DB 2107 AACAACTGGGGCATCGACCCAAAGCCATCGGGTCAAAATCTTCAACATCCAGGTCAAG 2048  
QY 323 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 342  
DB 2047 GAGGTACAGACGTGCAACGGCGAGACAACGGTGGCTAATAACCTTACCAGCACGGTTTCA 1988  
QY 343 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 362  
DB 1987 ATCTTTGGGACTCGTCTGACGAACCTGCGGTACGTGATGGATGGGGTCAAGAGGGCAGC 1928  
QY 363 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyr-----Leu 379

DB 1927 CTGCTCTCTTTTCCCAACGAGCTCTTTATGGTGCCCCAGTACGGCTACTGTGGAGTGGTG 1868  
QY 380 ThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhe 399  
DB 1867 ACCGCAACACTTCGCAGCAACAGACTGACAGAAATGCCTTCTACTGCCTGGAGTACTTT 1808  
QY 400 ProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluVal 419  
DB 1807 CCTTCGAGATGCTCGGACTTGGCAACAACCTTTGAAATTACGTACAGTTTTTGAGAAGGTG 1748  
QY 420 ProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIle 439  
DB 1747 CCTTCCACTCGATGTACGCGCACAGCCAGAGCCTGGACCGGCTGATGAACCCCTCTCATC 1688  
QY 440 AspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsn-----GlnSerGlySer 455  
DB 1687 GACCAGTACCTGTGGGACTGCAATCGACCACCACCGGMACCACCCTGAATGCCGGGACT 1628  
QY 456 AlaGlnAsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnPro 475  
DB 1627 GCCACCACCAAC-----TTTACCAAGCTGCGGCCCTACCAACTTTTCCAACCTTAAA 1577  
QY 476 LysAsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAsp 495  
DB 1576 AAGAACTGGCTGCCCGGCCCTCAATCAAGCAGCAGGGCTTCTCAAAGACTGCC----- 1523  
QY 496 AsnAsnAsnSerAsnPheThrTrpThrGlyAlaSer-----LysTyr----- 509  
DB 1522 AATCAAAACTACAAGATCCCTGCCACCGGGTTCAGACAGTCTCATCAAAATACGAGACGCAC 1463  
QY 510 ---AsnLeuAsnGlyArgGluSerIleIleAsnProGlyThrAlaMetAlaSerHisLys 528  
DB 1462 AGCACTCTGGACGGAAGATGGAGTGCCCTGACCCCGGACCTCCAATGGCCACGGCTGGA 1403  
QY 529 AspAspGluAspLysPhePheProMetSerGlyValMetIlePheGlyLysGluSerAla 548  
DB 1402 CCTCGGACAGCAAG--TTCAAGCAACAGCCAGCTCATCTTTGCGGGGCTTAAACAGAAC 1346  
QY 549 GlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAspGluGluGluIleLysAla 568  
DB 1345 GGCAACACGGCCACCGTACCCGGGACTCTGATCTTCACTCTGAGGAGGAGCTGGCAGCC 1286  
QY 569 ThrAsnProValAlaThrGluArgPheGlyThrValAlaValAsnPheGlnSerSerSer 588  
DB 1285 ACCAACGCCACCGATACGGACATGTGGGGCAACCTACCTGGCGGTGACCAAGCAACAGC 1226  
QY 589 ThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeuProGlyMetValTrpGln 608  
DB 1225 AACCTGCCGACCGTGGACAGACTGACAGCCTTGGGAGCCGTGCCCTGGAATGGTCTGGCAA 1166  
QY 609 AspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIleProHisThrAspGlyHis 628  
DB 1165 AACAGAGACATTACTACCAGGGTCCCATTTGGGCCAAGATTCTCTATACCGATGGACAC 1106  
QY 629 PheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsnProProProGlnIleLeu 648  
DB 1105 TTTCAACCCCTCACCGCTGATTGGTGGGTTTGGGCTGAACACACCCCGCTCTCAAAATTTT 1046  
QY 649 IleLysAsnThrProValProAlaAsnProProAlaGluPheSerAlaThrLysPheAla 668  
DB 1045 ATCAAGAACAACCCCGGTACCTGCGAATCTCTGCAACGACCTTCAGCTCTACTCCGGTAAAC 986  
QY 669 SerPheIleThrGlnTyrSerThrGlyGlnValSerValGluIleGluTrpGluLeuGln 688  
DB 985 TCCTTCATTACTCAGTACAGCACTGGCCAGGTGTGGTGAGATTGACTGGGAGATCCAG 926  
QY 689 LysGluAsnSerLysArgTrpAsnProGluValGlnTyrThrSerAsnTyrAlaLysSer 708  
DB 925 AAGGAGCGGTCCAAACGCTGGAACCCCGAGGTCCAGTTTACCTCCAACCTACGGACAGCAA 866  
QY 709 AlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThrGluProArgProIleGly 728

Db 865 AACTCTCTGTGTGGGCTCCCGATGCGGCTGGGAAATACACTGAGCCTAGGCGTATCGGT 806

Qy 729 ThrArgTyrLeuThrArgProLeu 736

Db 805 ACCCGCTACCTCACCCACCACTG 782

RESULT 12

US-09-532-594B-1

; Sequence 1, Application US/09532594B

; Patent No. 6468524

; GENERAL INFORMATION:

; APPLICANT: Chorini, John A.

; APPLICANT: Kotin, Robert M.

; APPLICANT: Safer, Brian

; APPLICANT: Davidson, Beverly

; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF

; FILE REFERENCE: 14014.0252U2

; CURRENT APPLICATION NUMBER: US/09/532,594B

; CURRENT FILING DATE: 2000-03-22

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 4767

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =

; NAME/KEY: misc\_feature

; LOCATION: 3009

; OTHER INFORMATION: n = g, a, c or t (u)

; NAME/KEY: misc feature

; OTHER INFORMATION: AAV4 genome

US-09-532-594B-1

Alignment Scores:

|                        |           |               |      |
|------------------------|-----------|---------------|------|
| Pred. No.:             | 1.28e-229 | Length:       | 4767 |
| Score:                 | 2481.50   | Matches:      | 474  |
| Percent Similarity:    | 74.20%    | Conservative: | 81   |
| Best Local Similarity: | 63.37%    | Mismatches:   | 162  |
| Query Match:           | 62.21%    | Indels:       | 31   |
| DB:                    | 4         | Gaps:         | 10   |

US-09-807-802A-13 (1-736) x US-09-532-594B-1 (1-4767)

Qy 4 AspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArgGluTrpTrp 23

Db 2266 GACGGTTACCTTCCAGATTGGCTAGAGGACAACTCTCTGAAGCGGTTTCGAGAGTGGTGG 2325

Qy 24 AspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAspAspGlyArg 43

Db 2326 GCGTGCAACTGGAGCCCTTAACCCAGGCAATCAACAACTCAGGACAAACGCTCGG 2385

Qy 44 GlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAspLysGlyGlu 63

Db 2386 GGTCTTGTGCTTCGGGTTTACAAATACCTCGGACCCGCAACGGACTCGACAAAGGGGAA 2445

Qy 64 ProValAsnAlaAspAlaAlaLeuGluHisAspLysAlaTyrAspGlnGlnLeu 83

Db 2446 CCCGTCAACGCGGACGCGGCGAGCCCTCGAGCACGACAAAGGCCTACGACCAGCAGCTC 2505

Qy 84 LysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPheGlnGluArg 103

Db 2506 AAGCCCGGTGACAAACCCCTACCTCAAGTACACCAACCGCCGACGCGGAGTTCCAGCAGCG 2565

Qy 104 LeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGlnAlaLysLys 123

Db 2566 CTTCAGGGCGACACACCGTTTGGGGGCAACCTCGGCAGAGCAGTCTTCCAGGGCCAAAAG 2625

Qy 124 ArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaProGlyLysLys 143

Db 2626 AGGGTTCTTGAACCTCTTGGTCTGGTTGAGCAAGCGGGTGAGACGGGCTCTCTGGAAAGAA 2685

Qy 144 ArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGlyLysThrGly 163

Db 2686 AGACCGTTGATTGAATCCCCCAGCAGCCCGACTCTCCACGGGTATCGCAAAAAGGC 2745

Qy 164 GlnGlnProAlaLysLysArgLeuAsnPhe---GlyGlnThrGlyAspSerGluSerVal 182

Db 2746 AAGCAGCCGGCTAAAAAAGAGTCGTTTTTCGAAGACGAAACTGGAGCAGCGACGGACCC 2805

Qy 183 ProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyProThrThr 202

Db 2806 CCTGAGGGATCAACTTCGGGA-----GCCATGTCTGTGATGACAGTGAG 2847

Qy 203 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 222

Db 2848 ATGCGTGCAGCAGCTGGCGGAGCTGCAGTCGAGGGSGGACNAGGTGCCGATGGAGTGGGT 2907

Qy 223 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 242

Db 2908 AATGCCTCGGGTGATTGGCATTTGCCATTCCACCTGGTCTGAGGGCCACGTCACGACCAC 2967

Qy 243 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 262

Db 2968 AGCACCAGAACTGGGTCTTGCCCACTTACAAACCAACCACTTNTACAAGCGACTCGGAGAG 3027

Qy 263 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 282

Db 3028 -----AGCCTGCAGTCCAACACCTACAACGGATTCTCCACCCCTGGGGATAC 3075

Qy 283 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 302

Db 3076 TTTGACTTCAACCGCTTCCACTGCCACTTCTCACCACTGACTGGCAGCGACTCATCAAC 3135

Qy 303 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 322

Db 3136 AACAACTGGGCGATGCGACCCCAAGCCATGCGGTCAAAATCTTCAACATCCAGGTCAAG 3195

Qy 323 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 342

Db 3196 GAGGTCACGACGTCGAAACGGCGAGACAACGGTGTCTAATAACCTTACCAGCACGGTTCAG 3255

Qy 343 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 362

Db 3256 ATCTTTGCGGACTCGTCGTACGAACCTGCCGTACGTGATGGATGCGGGTCAAGAGGGCAGC 3315

Qy 363 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyr-----Leu 379

Db 3316 CTGCCTCCTTTTCCCAACGACGCTCTTTATGGTGCCCCAGTACGGCTACTGTGGACTGGTG 3375

Qy 380 ThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhe 399

Db 3376 ACCGGCAACACTTCGCGAGCAACAGACTGACAGAAATGCCTTCTACTGCCTGGAGTACTTT 3435

Qy 400 ProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluVal 419

Db 3436 CCTTCGAGATGCTGCGGACTGGCAACAACCTTTTGAAATTACGTACAGTTTTTGAGAAGGTG 3495

Qy 420 ProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIle 439

Db 3496 CCTTTCACACTCGATGTACGCGCACAGCCAGAGCCTGACCGGCTGATGAACCCCTCTCATC 3555

Qy 440 AspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsn-----GlnSerGlySer 455

Db 3556 GACCAGTACCTGTGGGGACTGCAATCGACCACCCACCGGAACCACTGAATGCCGGGACT 3615

Qy 456 AlaGlnAsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnPro 475

Db 3616 GCCACCACCAAC-----TTTACCAAGCTGGGCGCTACCAACTTTTCCAACTTTAAA 3666

Qy 476 LysAsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAsp 495

Db 3667 AAGAACTGGCTGCCCGGCGCTTCAATCAATCAAGCAGCAGGGCTTCTCAAAGACTGCC----- 3720

Qy 496 AsnAsnAsnSerAsnPheThrTrpThrGlyAlaSer-----LysTyr----- 509



```
Db 3721 AATCAAACTACAAGATCCCTGCCACCGGGTCAGACAGTCTCATCAATACGAGACGCAC 3780
Qy 510 ---AsnLeuAsnGlyArgGluSerIleIleAsnProGlyThrAlaMetAlaSerHisLys 528
Db 3781 AGCACTCTGGACGGAAGATGGAGTGCCTGACCCCGGACCTCCAATGGCCACGGCTGGA 3840
Qy 529 AspAspGluAspLysPhePheProMetSerGlyValMetIlePheGlyLysGluSerAla 548
Db 3841 CCTGGGACAGCAAG--TTCAGCAACAGCCAGCTCATCTTTGCGGGCCCTAAACAGAAC 3897
Qy 549 GlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAspGluGluIleLysAla 568
Db 3898 GGCAACACGGCCACCGTACCCGGGACTCTGATCTTCACCTCTGAGGAGGAGCTGGCAGCC 3957
Qy 569 ThrAsnProValAlaThrGluArgPheGlyThrValAlaValAsnPheGlnSerSerSer 588
Db 3958 ACCAAGCCACCGCATACGGACATGTGGGGCAACCTACCTGGCGGTGACCAGAGCAACAGC 4017
Qy 589 ThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeuProGlyMetValTrpGln 608
Db 4018 AACCTGCCGACCGTGGACAGACTGACAGCCTTGGGAGCGCGTCTGGAATGGTCTGCA 4077
Qy 609 AspArgAspValTyrLeuGlnGlyProIleTyrAlaLysIleProHisThrAspGlyHis 628
Db 4078 AACAGAGACATTTACTACCAGGGTCCCATTTGGGCCAAGATTCCTCATACCGATGCAC 4137
Qy 629 PheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsnProProProGlnIleLeu 648
Db 4138 TTTACCCCTCACCGCTGATTGGTGGGTTGGGCTGAAACACCCGCTCCTCAAAATTTT 4197
Qy 649 IleLysAsnThrProValProAlaAsnProProAlaGluPheSerAlaThrLysPheAla 668
Db 4198 ATCAAGAACACCCCGGTACTCTGCGAATCCTGCAACGACCTTCAGCTCTACTCCGGTAA 4257
Qy 669 SerPheIleThrGlnTyrSerThrGlyGlnValSerValGluIleGluTrpGluLeuGln 688
Db 4258 TCCTTCATTACTCAGTACAGCACTGGCCAGGTGTCGGTGCAGATTGACTGGGAGATCCAG 4317
Qy 689 LysGluAsnSerLysArgTyrAsnProGluValGlnTyrThrSerAsnTyrAlaLysSer 708
Db 4318 AAGGAGCGGTCCAAACGCTGGAACCCGAGGTCCAGTTTACCTCCAACTACGGACAGCAA 4377
Qy 709 AlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThrGluProArgProIleGly 728
Db 4378 AACTCTCTGTGTGGGCTCCCGATGGCGCTGGGAAATACACTGAGCCTAGGGCTATCGGT 4437
Qy 729 ThrArgTyrLeuThrArgProLeu 736
Db 4438 ACCCGCTACTCACCCACCACCTG 4461
```

RESULT 13

```
US-09-532-594B-17
; Sequence 17, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 1800
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: Synthetic construct
```

```
; NAME/KEY: misc_feature
; LOCATION: 342
; OTHER INFORMATION: n = g, a, c or t(u)
; NAME/KEY: misc_feature
; OTHER INFORMATION: AAV4 capsid protein VP2 gene
US-09-532-594B-17

Alignment Scores:
Pred. No.: 2,72e-167 Length: 1800
Score: 1830.50 Matches: 355
Percent Similarity: 69.87% Conservative: 74
Best Local Similarity: 57.82% Mismatches: 154
Query Match: 45.89% Indels: 31
DB: 4 Gaps: 10

US-09-807-802A-13 (1-736) x US-09-532-594B-17 (1-1800)
Qy 138 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 157
Db 1 ACGGCTCCTGGAAGAAGAGACCGTTGATTGAATCCCCCAGCAGCCGACTCCTCCACG 60
Qy 158 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPhe---GlyGlnThr 176
Db 61 GGTATCGGCAAAAAGCAAGCAGCCGGCTAAAAGAGCTCGTTTTCGAAGACGAAACT 120
Qy 177 GlyAspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAla 196
Db 121 GGAGCAGCGCAGCAGCCCTCTGAGGGATCAACTTCCGGA-----GCC 162
Qy 197 AlaValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGlu 216
Db 163 ATGTCTGATGACAGTGAGATCGTGACGAGCTGGCGGAGCTGCAGTCGAGGSGGACAA 222
Qy 217 GlyAlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGly 236
Db 223 GGTGCCGATGGAGTGGGTAATGCCCTCGGGTGATTGGCATTCGATTCACCTGCTCTGAG 282
Qy 237 AspArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeu 256
Db 283 GGCCACGTCACGACCAACAGCAGCAGAACCTGGGTCTTGCCACCTACACCAACACCTN 342
Qy 257 TyrLysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyr 276
Db 343 TACAAGCGACTCGGAGAG-----AGCCTGCAGTCCCAACACCTACACCGGATTC 390
Qy 277 SerThrProTyrGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAsp 296
Db 391 TCCACCCCTGGGGATACTTTTGACTTCAACCGCTTCCACTGCCACTTCTCACCACGTGAC 450
Qy 297 TrpGlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeu 316
Db 451 TGGCAGCGACTCATCAACAACAACCTGGGGGATGCGACCCCAAGCCATCGGGGTCAAAATC 510
Qy 317 PheAsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsn 336
Db 511 TTCAACATCCAGGTCAAGGAGGTTCACGACGTGCAACGGCGGAGACAAACGGTGGCTAATAAC 570
Qy 337 LeuThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGly 356
Db 571 CTTACCAGCACCGTTTCAGATCTTTGCGGACTCTGCTGTAAGAACTGCGGTACGTGATGAT 630
Qy 357 SerAlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyr 376
Db 631 GCGGGTCAAGAGGGGCGCCTGCCTCTTTCACCAACGACGTCTTTATGTTGTCGCCAGTAC 690
Qy 377 GlyTyr-----LeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPhe 393
Db 691 GGCTACTGTGGACTGGTGACCGGCAACACTTCGAGCAACAGACTGACAGAAATGCCTTC 750
Qy 394 TyrCysLeuGluTyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSer 413
Db 751 TACTGCCTGGAGTACTTTCTTCGAGATGCTCGGAGACTGGCAACAACTTTGAAATTACG 810
```

```
Qy 414 TyrThrPheGluGluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArg 433
Db 811 TACAGTTTGGAGAGGTGCTTCCACTCGATGTACGGGCACAGCCAGAGCCTGGACCGG 870
Qy 434 LeuMetAsnProLeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsn----- 451
Db 871 CTGATGAACCTCTCATCGACAGTACCTGTGGGACTGGAATCGACACCACCGGAACC 930
Qy 452 -----GlnSerGlySerAlaGlnAsnLysAspLeuLeuPheSerArgGlySerProAla 469
Db 931 ACCCTGAATCGCGGACTGCCACCAAC-----TTACCAAGCTGCGGCCTACC 981
Qy 470 GlyMetSerValGlnProLysAsnTrpLeuProGlyProCysTyrArgGlnGlnArgVal 489
Db 982 AACTTTTCCAACTTTAAAGAACTGGCTGCCCGGCCCTTCAATCAAGCAGCAGGGCTTC 1041
Qy 490 SerLysThrLysThrAspAsnAsnAsnSerAsnPheThrTrpThrGlyAlaSer----- 507
Db 1042 TCAAAGACTGCC-----AATCAAAACTACAAGATCCCTGCCACCGGTTCAGACAGTCTC 1095
Qy 508 ---LysTyr-----AsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 522
Db 1096 ATCAATACGACGACGACAGCACTCTGGACGGAAGATGGAGTGCCTTGACCCCGGACCT 1155
Qy 523 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 542
Db 1156 CCAATGGCCACGGCTGGACCTGCGGACAGCAAG---TTCAGCAACAGCCAGCTCATCTTT 1212
Qy 543 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 562
Db 1213 GCGGGCCCTAAACAGAACGCAACACGCGCCACCGTACCCGGGACTCTGATCTTCACCTCT 1272
Qy 563 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 582
Db 1273 GAGGAGGAGTGGCAGCCACCAACGCCACCGATACGGACATGTGGGCAACCTACCTGSC 1332
Qy 583 AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 602
Db 1333 GGTGACCAGACGACAGCAACCTGCCGACCGTGGACAGACTGACAGCCTTGGGAGCCCTG 1392
Qy 603 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 622
Db 1393 CCTGGAATGGTCTGGCAAAACAGAGACATTTACTACGAGGTCCCATTTGGGCCAAGATT 1452
Qy 623 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 642
Db 1453 CCTCATACCGATGGACACTTTTACCCCTTACCCGCTGATTGGTGGGTTGGCTGAAACAC 1512
Qy 643 ProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 662
Db 1513 CCGCCTCTCTCAAAATTTTATCAAGAACACCCCGGTACCTCGGAATCCTGCAACGACCTTC 1572
Qy 663 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 682
Db 1573 AGCTCTACTCCGCTAAACTCCTTCTATTACTCAGTACAGCACTGGCCAGGTGTCGGTGCAG 1632
Qy 683 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 702
Db 1633 ATTGACTGGGAGATCCAGAAGGAGCGGTCCAAACGCTGGAACCCCGAGGTCCAGTTTACC 1692
Qy 703 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 722
Db 1693 TCCAACCTACGACAGCAAAACTCTCTGTGTGGGCTCCCGATGCGGCTGGGAAATACACT 1752
Qy 723 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736
Db 1753 GAGCCTAGGGCTATCGGTACCCGCTACCTCACCCACACCTG 1794
```

RESULT 14

US-09-532-594B-19

; Sequence 19, Application US/09532594B

; Patent No. 6468524

; GENERAL INFORMATION:

```
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 1617
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: synthetic construct
; NAME/KEY: misc_feature
; LOCATION: 162
; OTHER INFORMATION: n = g, a, c or t(u)
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 capsid protein VP3 gene
US-09-532-594B-19
```

```
Alignment Scores:
Pred. No.: 8.17e-154 Length: 1617
Score: 1690.50 Matches: 325
Percent Similarity: 71.30% Conservative: 65
Best Local Similarity: 59.41% Mismatches: 132
Query Match: 42.38% Indels: 25
DB: 4 Gaps: 9
```

US-09-807-802A-13 (1-736) x US-09-532-594B-19 (1-1617)

```
Qy 204 AlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGlyAsn 223
Db 7 GCAGCAGCTGGCGAGCTGCAGTC---GAGGSGGACAAAGGTGCCGATGGAGTGGTAAAT 63
Qy 224 AlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThrSer 243
Db 64 GCCTCGGGTGATGGCATTTGCGATTCCACCTGTCTGAGGGCCACGTCCAGCACCAGC 123
Qy 244 ThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSerAla 263
Db 124 ACCAGAACCTGGGTCTTGGCCACCTTACCAACACCACTTACCAAGCGACTCGGAGAG--- 180
Qy 264 SerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyrPhe 283
Db 181 -----AGCCTGCAGTCCAAACACCTACACGGATTCTCCACCCCTGGGATACTTT 231
Qy 284 AspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsnAsn 303
Db 232 GACTTCAACCGCTTCCACTGCCACTTCTCACCAGCTGACTGGCAGCGACTCATCAACAAC 291
Qy 304 AsnTrpGlyPheArgProLysArgLeuAsnAsnPheLysLeuPheAsnIleGlnValLysGlu 323
Db 292 AACTGGGCGCATCGACCCAAAGCCATGCGGGTCAAAATCTTCAACATCCAGGTCAAGGAG 351
Qy 324 ValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGlnVal 343
Db 352 GTCACGACGTGCAACGCGGAGACACCGTGGCTTAATAACCTTACCAGCACGGTTCAGATC 411
Qy 344 PheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCysLeu 363
Db 412 TTTGCGGACTCGTCTGACGAACCTGCCGTACGTGATGGATGCGGGTCAAGAGGCGACCTG 471
Qy 364 ProProPheProAlaAspValPheMetIleProGlnTyrGlyTyr-----LeuThr 380
Db 472 CCTCCTTTTCCCAACGACGCTCTTTATGGTGGCCCGCCAGTACGGCTACTGTGGAGCTGGTGACC 531
Qy 381 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 400
Db 532 GGCAACACACTTCGACAGCAACAGACTGACAGAAATGCTTCTACTGCTGAGTACTTCTCCT 591
```

```
Qy 401 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluValPro 420
Db 592 TCGCAGATGCTGGGACTGGCAACAACCTTTGAATACGTACAGTTTGAAGGTGCCT 651
Qy 421 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 440
Db 652 TTCCACTCGATGACGGCACAGCCAGAGCCTGGACCGGCTGATGAACCTCTCATCGAC 711
Qy 441 GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsn-----GlnSerGlySerAla 456
Db 712 CAGTACCTGTGGGAGCTGCAATCGACCAACCACCGGAACCCCTGAATGCCGGGACTGCC 771
Qy 457 GlnAsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLys 476
Db 772 ACCACCAAC-----TTTACCAAGCTGGGGCTACCAACTTTTCCAACTTTAAAAAG 822
Qy 477 AsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsn 496
Db 823 AACTGGCTGCCCGGCTTCAATCAAGCAGCAGGGCTTCTCAAGACTGCC-----AAT 876
Qy 497 AsnAsnSerAsnPheThrTrpThrGlyAlaSer-----LysTyr----- 509
Db 877 CAAAACTACAAGATCCCTGCCACCGGTCAGACACTCTCATCAATACGAGACGCACAGC 936
Qy 510 AsnLeuAsnGlyArgGluSerIleIleAsnProGlyThrAlaMetAlaSerHisLysAsp 529
Db 937 ACTCTGGACGGAAGATGGAGTGCCTGACCCCGGACCTCCAATGGCCACGGCTGGACCT 996
Qy 530 AspGluAspLysPhePheProMetSerGlyValMetIlePheGlyLysGluSerAlaGly 549
Db 997 GCGGACAGCAAG---TTCAGCAACAGCCAGCTCATCTTTGGGGGCTTAAACAGAACGGC 1053
Qy 550 AlaSerAsnThrAlaLeuAspAsnValMetIleThrAspGluGluIleLysAlaThr 569
Db 1054 AACACGGCCACCGTACCCGGGACTCTGATCTTACCTCTGAGGAGGAGCTGGGAGCCACC 1113
Qy 570 AsnProValAlaThrGluArgPheGlyThrValAlaValAsnPheGlnSerSerThr 589
Db 1114 AACGCCACCGATACGGACATGTGGGGCAACCTACCTGGCGGTGACCAGAGCAACAGCAAC 1173
Qy 590 AspProAlaThrGlyAspValHisAlaMetGlyAlaLeuProGlyMetValTrpGlnAsp 609
Db 1174 CTGCCGACCGTGGACAGACTGACAGCCTTGGGAGCCCTGGGATGGTCTGGCAAAAC 1233
Qy 610 ArgAspValTyrLeuGlnGlyProIleTrpAlaLysIleProHisThrAspGlyHisPhe 629
Db 1234 AGAGACATTTACTACGAGGTCCCATTTGGGCCAAGATTCTCATACCGATGGACACTTT 1293
Qy 630 HisProSerProLeuMetGlyGlyPheGlyLeuLysAsnProProProGlnIleLeuIle 649
Db 1294 CACCCCTCACCGCTGATTGGTGGGTTGGGCTGAAACACCCGCTCTCTCAAATTTTATC 1353
Qy 650 LysAsnThrProValProAlaAsnProProAlaGluPheSerAlaThrLysPheAlaSer 669
Db 1354 AAGAACACCCCGGTACCTGCGAATCCTGCAACGACCTTCAGCTCTACTCCGGTAAACTCC 1413
Qy 670 PheIleThrGlnTyrSerThrGlyGlnValSerValGluIleGluTrpGluGlnLys 689
Db 1414 TTCAATTACTCAGTACAGCACTGGCCAGGTGTCGGTGCAGATTGACTGGGAGATCCAGAAG 1473
Qy 690 GluAsnSerLysArgTyrAsnProGluValGlnTyrThrSerAsnTyrAlaLysSerAla 709
Db 1474 GAGCGGTCCAAACGCTGGAACCCCGAGGTCCAGTTTACTCTCAACTACGGACAGCAAAAC 1533
Qy 710 AsnValAspPheThrValAspAsnAsnGlyLeuTyrThrGluProArgProIleGlyThr 729
Db 1534 TCTCTGTTGTGGGCTCCCGATGGCGCTGGGAAATACACTGAGCCTAGGGCTATCGGTACC 1593
Qy 730 ArgTyrLeuThrArgProLeu 736
Db 1594 CGCTACCTCACCCACCACCTG 1614
```

```
RESULT 15
US-09-438-268-3
; Sequence 3, Application US/09438268
; Patent No. 6491907
; GENERAL INFORMATION:
; APPLICANT: Rabinowitz, Joseph E.
; APPLICANT: Samulski, Richard J
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 5470-186
; CURRENT APPLICATION NUMBER: US/09/438,268
; EARLIER APPLICATION NUMBER: 60/107,840
; EARLIER FILING DATE: 1998-11-10
; EARLIER APPLICATION NUMBER: 60/123,651
; EARLIER FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2271
; TYPE: DNA
; ORGANISM: Virus
US-09-438-268-3

Alignment Scores:
Pred. No.: 2,07e-126 Length: 2271
Score: 1410.00 Matches: 312
Percent Similarity: 55.75% Conservative: 95
Best local Similarity: 42.74% Mismatches: 269
Query Match: 35.35% Indels: 54
DB: 4 Gaps: 16
```

US-09-807-802A-13 (1-736) x US-09-438-268-3 (1-2271)

```
Qy 1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20
Db 1 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACACTCTCTCTGAAGGAATAAGA 60
Qy 21 GluTrpTrpAspLeuLysProGlyValaProLysProLysAlaAsnGlnLysGlnAsp 40
Db 61 CAGTGGTGGAAAGCTCAAACCTGGCCACCACCACCAAGCCCGCAGAGCGGCGGACATAAGGAC 120
Qy 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp 60
Db 121 GACAGCAGGGGTCTTGTGCTTCTCGGTACAAGTACCTCGGACCTTCAACGGACTCGAC 180
Qy 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80
Db 181 AAGGAGAGCCGGTCAACGAGGACAGCCCGCGGCTCGAGCAGCAACAAAGCCTACGAC 240
Qy 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100
Db 241 CGGCAGCTCGACAGCGGAGACAACCCGTACCTCAAGTACAACCAACCGCGGAGTTT 300
Qy 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120
Db 301 CAGGAGCGCCTTAAAGAAAGATACGTCTTTGGGGGCAACCTCGGACGAGCAGTCTTCCAG 360
Qy 121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro 140
Db 361 GCGAAAAAGAGGGTCTTTGAACCTCTGGGCTGGTTGAGGAACCTGTAAAGACGGCTCCG 420
Qy 141 GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly 160
Db 421 GGAAAAAAGAGSCGGTAGAGCACTCTCTCTGTGGAGCCAGACTCCTCCTCGGAAACCGGA 480
Qy 161 LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180
Db 481 AAGCGGGGCCAGCAGCCTGCAAGAAAAAGATTGAATTTTGGTCAGACTGGAGACGCAGAC 540
Qy 181 SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro 200
```



Db 541 TCAGTACTGACCCCGCCCTCTCGGACAGCCACCAGCAGCCCGCCCTCTGGTCTGGGAAC 600  
Qy 201 ThrThrMetAlaSerGlyGlyAlaPrometAlaAspAsnAsnGluGlyAlaAspGly 220  
Db 601 AATACGATGACTTCAGTTAATCTGCAGAA---GCCAGCACTGGTGCAGGAGGGGGGGC 657  
Qy 221 ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle 240  
Db 658 AGTAATCTGTCAAAAGCATGTGGAGTGAGGGGCCACTTTAGTGTAACTCTGTAACT 717  
Qy 241 ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle 260  
Db 718 TGTACATTTCCAGACAGTTTTAAATTCATATGACCCAGACCATTAATAAGGTGTTT 777  
Qy 261 SerSerAlaSerThrGlyAlaSerAsnAsp----- 270  
Db 778 TCTCCCGCAGCGAGTAGTGCCACAATGCCAGTGGAAGGAGGCAAAAGGTTTGCACCATC 837  
Qy 271 AsnHisTyrPheGlyTyrSerThrProTrpGlyTyrPheAspPheAsnArgPheHisCys 290  
Db 838 AGTCCCAATAATGGGATACTCAACCCCATGGAGATATTTAGATTATTAATGCTTTAAATTTA 897  
Qy 291 HisPheSerProArgAspTrpGlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLys 310  
Db 898 TTTTTCACCTTTAGAGTTTCAGACTTAATTGAAATTTATGGAAGTATAGTCTCTGAT 957  
Qy 311 ArgLeuAsnPheLysLeuPheAsnIleGlnValLysGluValThr-----ThrAsnAsp 328  
Db 958 GCTTTAACTGTAAACCATATCAGAAATTGCTGTTAAGGATGTTACAGACAAAACCTGGAGGG 1017  
Qy 329 GlyValThrThrIleAlaAsnAsnLeuThrSerThrValGlnValPheSerAspSerGlu 348  
Db 1018 GGGGTA---CAGGTTACTGACAGCACTACAGGGCGCCTATGCGATGTTAGTAGACCATGAA 1074  
Qy 349 TyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCysLeuProPropheProAla 368  
Db 1075 TACAAGTACCATATGTGTTAGGGCAAGGTGAGGATACCTTTAGCCCCCAGAACTTCTCTATT 1134  
Qy 369 AspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsnAsn---GlySerGlnAla 387  
Db 1135 TGGGTATACCTTTCCCGCTCAATATGCTTACTTAACAGTAGGAGATGTTAACACACAAGGA 1194  
Qy 388 ValGly-----ArgSerSerPheTyrCysLeuGluTyr 398  
Db 1195 ATTTCTGGAGACAGCAAAATAATTAGCAAGTGAAGAAATCAGCATTTTATGTTTGGAAACAC 1254  
Qy 399 PheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGlu 418  
Db 1255 AGTTCTTTTTCAGCTTTTAGGTACAGGAGGTACAGCAACTATGCTTATAAGTTTCTCTCCA 1314  
Qy 419 ValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeu 438  
Db 1315 GTGCCCCCAGAAAATTTAGAGGGCTGCAGTCAACACACTTTTATGAATGTACAATCCCTTA 1374  
Qy 439 IleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsn 458  
Db 1375 -----TACGGATCCCGCTTAGGGGTTCTCTGACACATTAGGAGGTGACCCCAAAATTT 1425  
Qy 459 LysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrp 478  
Db 1426 AGATCTTTA-----ACACATGAAGACCATGCAATTCAGCCCCCAAACTTC 1470  
Qy 479 LeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsn 498  
Db 1471 ATGCCAGGGCCACTAGTAAACTCAGTGTCTACAAAGGAGGGAGACAGCTCTAATACTGGA 1530  
Qy 499 SerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIle 518  
Db 1531 GCTGGAAAAGCCCTTAACAGGCCCTTAGCACAGGTACCTCTCAAAACACTAGATAATATCCTTA 1590  
Qy 519 AsnProGly---ThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 537  
Db 1591 CGCCCTGGGCCAGTGTCTCAGCCCATACCCACTGGGACACAGATAAATATGTACACAGGA 1650

\* . . \*

Qy 538 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 557  
Db 1651 ATAAATGCCATTCTCTCATGGTCAGACCCTTATGGT-----AACGCTGAAGACAAA 1701  
Qy 558 Val-----MetIleThrAspGluGluGluIleLysAlaThrAsn 570  
Db 1702 GAGTATCAGCAAGGAGTGGTAGATTTCCAAATGAAAAAGAACAGCTAAAAACAGTTACAG 1761  
Qy 571 ProValAlaThrGluArgPheGlyThrValAlaValAsnPheGlnSerSerThrAsp 590  
Db 1762 GGTTTAAACATGCACACCTAC-----TTTCCCAATAAAAGGAACCCAG 1803  
Qy 591 ProAlaThrGlyAspValHisAlaMetGlyAlaLeuProGlyMetValTrpGlnAspArg 610  
Db 1804 CAATATACAGATCAAAATTGAG---CGCCCCCTAATGGTGGTTCGTATGGAACAGAAGA 1860  
Qy 611 AspValTyrLeuGlnGlyProIleTrpAlaLysIleProHisThrAspGlyHisPheHis 630  
Db 1861 GCCCTTCACTATGAAAGCCAGCTGTGGAGTAAATTCCAAATTTAGATGACAGTTTTAA 1920  
Qy 631 ProSer---ProLeuMetGlyGlyPheGlyLeuLysAsnProProGlnIleLeuIle 649  
Db 1921 ACTCAGTTTGCAGCTTAGGAGGATGGGGTTTGCATCAGCCACCTCCTCAAAATATTTTAA 1980  
Qy 650 LysAsnThrProValProAlaAsnProProAlaGluPheSerAlaThrLysPheAlaSer 669  
Db 1981 AAA-----ATATTACCACAAAAGTGGCCAAATTGGAGGTATTTAAATCAATGGGAATTACT 2034  
Qy 670 PheIleThrGlnTyrSerThrGlyGlnValSerValGluIleGluTrpGluLeu---Gln 688  
Db 2035 ACCTTAGTTTCAGTATGCCGTGGGAATTATGACAGTAACATATGACATTTAAATTTGGGGCCC 2094  
Qy 689 LysGluAsnSerLysArgTrpAsnProGlu 698  
Db 2095 CGTAAAGCTACGGGACGGTGGAAATCCTCAA 2124

Search completed: February 15, 2004, 01:34:58  
Job time : 286.687 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 14, 2004, 16:55:45 ; Search time 684.807 Seconds  
(without alignments)  
3959.005 Million cell updates/sec

Title: US-09-807-802A-13  
Perfect score: 3989  
Sequence: 1 MAADGYLPDWLEDNLSEGR.....NNGLYTEPRPIGTRYLTRPL 736

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2449703 seqs, 1841816367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09807802/runat\_11022004\_175610\_15987/app\_query.fasta\_1.2389  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09807802 @CGN 1 1 1100 @runat\_11022004\_175610\_15987  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq2:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB ID | Description |
|------------|-------|-------|--------|-------|-------------|
| -----      |       |       |        |       |             |

|    |        |       |      |    |                   |                   |
|----|--------|-------|------|----|-------------------|-------------------|
| 1  | 3989   | 100.0 | 4718 | 13 | US-10-291-583-6   | Sequence 6, Appli |
| 2  | 3984   | 99.9  | 7447 | 13 | US-10-216-870-11  | Sequence 11, Appl |
| 3  | 3494   | 87.6  | 4726 | 13 | US-10-291-583-8   | Sequence 8, Appl  |
| 4  | 3473   | 87.1  | 3122 | 13 | US-10-291-583-45  | Sequence 45, Appl |
| 5  | 3464   | 86.8  | 3117 | 13 | US-10-291-583-43  | Sequence 43, Appl |
| 6  | 3460   | 86.7  | 3121 | 13 | US-10-291-583-44  | Sequence 44, Appl |
| 7  | 3457   | 86.7  | 3122 | 13 | US-10-291-583-42  | Sequence 42, Appl |
| 8  | 3453   | 86.6  | 3128 | 13 | US-10-291-583-27  | Sequence 27, Appl |
| 9  | 3451.5 | 86.5  | 3075 | 13 | US-10-291-583-26  | Sequence 26, Appl |
| 10 | 3448   | 86.4  | 3129 | 13 | US-10-291-583-59  | Sequence 59, Appl |
| 11 | 3446.5 | 86.4  | 3142 | 13 | US-10-291-583-25  | Sequence 25, Appl |
| 12 | 3445   | 86.4  | 3123 | 13 | US-10-291-583-41  | Sequence 41, Appl |
| 13 | 3442.5 | 86.3  | 4721 | 13 | US-10-291-583-1   | Sequence 1, Appli |
| 14 | 3441   | 86.3  | 3128 | 13 | US-10-291-583-47  | Sequence 47, Appl |
| 15 | 3440   | 86.2  | 3128 | 13 | US-10-291-583-28  | Sequence 28, Appl |
| 16 | 3439   | 86.2  | 8179 | 15 | US-10-205-942-5   | Sequence 5, Appli |
| 17 | 3438   | 86.2  | 3084 | 13 | US-10-291-583-39  | Sequence 39, Appl |
| 18 | 3434   | 86.1  | 3098 | 13 | US-10-291-583-37  | Sequence 37, Appl |
| 19 | 3432   | 86.0  | 3197 | 13 | US-10-291-583-29  | Sequence 29, Appl |
| 20 | 3429   | 86.0  | 3121 | 13 | US-10-291-583-11  | Sequence 11, Appl |
| 21 | 3424.5 | 85.8  | 3127 | 13 | US-10-291-583-15  | Sequence 15, Appl |
| 22 | 3423.5 | 85.8  | 2205 | 13 | US-10-291-583-120 | Sequence 120, App |
| 23 | 3423.5 | 85.8  | 4385 | 13 | US-10-291-583-5   | Sequence 5, Appli |
| 24 | 3422.5 | 85.8  | 3095 | 13 | US-10-291-583-24  | Sequence 24, Appl |
| 25 | 3422   | 85.8  | 3121 | 13 | US-10-291-583-13  | Sequence 13, Appl |
| 26 | 3420.5 | 85.7  | 3113 | 13 | US-10-291-583-31  | Sequence 31, Appl |
| 27 | 3420   | 85.7  | 3098 | 13 | US-10-291-583-9   | Sequence 9, Appli |
| 28 | 3417.5 | 85.7  | 3095 | 13 | US-10-291-583-23  | Sequence 23, Appl |
| 29 | 3415.5 | 85.6  | 3122 | 13 | US-10-291-583-56  | Sequence 56, Appl |
| 30 | 3415   | 85.6  | 2489 | 13 | US-10-291-583-35  | Sequence 35, Appl |
| 31 | 3414.5 | 85.6  | 3094 | 13 | US-10-291-583-22  | Sequence 22, Appl |
| 32 | 3414.5 | 85.6  | 3106 | 13 | US-10-291-583-34  | Sequence 34, Appl |
| 33 | 3413.5 | 85.6  | 3123 | 13 | US-10-291-583-54  | Sequence 54, Appl |
| 34 | 3413.5 | 85.6  | 3276 | 13 | US-10-291-583-38  | Sequence 38, Appl |
| 35 | 3413   | 85.6  | 2495 | 13 | US-10-291-583-36  | Sequence 36, Appl |
| 36 | 3411.5 | 85.5  | 3113 | 13 | US-10-291-583-55  | Sequence 55, Appl |
| 37 | 3408.5 | 85.4  | 2504 | 13 | US-10-291-583-33  | Sequence 33, Appl |
| 38 | 3405   | 85.4  | 4393 | 12 | US-10-423-704A-1  | Sequence 1, Appli |
| 39 | 3405   | 85.4  | 4393 | 13 | US-10-291-583-4   | Sequence 4, Appli |
| 40 | 3404   | 85.3  | 3098 | 13 | US-10-291-583-10  | Sequence 10, Appl |
| 41 | 3402.5 | 85.3  | 4679 | 9  | US-09-804-898-1   | Sequence 1, Appli |
| 42 | 3402.5 | 85.3  | 4679 | 9  | US-09-945-681-10  | Sequence 10, Appl |
| 43 | 3402.5 | 85.3  | 4679 | 13 | US-10-136-819-6   | Sequence 6, Appli |
| 44 | 3402.5 | 85.3  | 4679 | 14 | US-10-038-972A-12 | Sequence 12, Appl |
| 45 | 3402   | 85.3  | 3106 | 13 | US-10-291-583-18  | Sequence 18, Appl |

ALIGNMENTS

RESULT 1  
US-10-291-583-6  
; Sequence 6, Application US/10291583  
; Publication No. US20030138772A1  
; GENERAL INFORMATION:  
; APPLICANT: Gao, Guangping  
; APPLICANT: Wilson, James M.  
; APPLICANT: Alvira, Mauricio  
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus  
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1el Sequences Identif  
; FILE REFERENCE: UPN-02735USA  
; CURRENT APPLICATION NUMBER: US/10/291,583  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: US 60/350,607  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/341,117  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: US 60/377,066  
; PRIOR FILING DATE: 2002-05-01  
; PRIOR APPLICATION NUMBER: US 60/386,675  
; PRIOR FILING DATE: 2002-06-05  
; NUMBER OF SEQ ID NOS: 120  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6

```
; LENGTH: 4718.
; TYPE: DNA
; ORGANISM: adeno-associated virus serotype 1
US-10-291-583-6

Alignment Scores:
Pred. No.:      0      Length:      4718
Score:          3989.00  Matches:      736
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:      100.00%  Indels: 0
DB:              13      Gaps: 0

US-09-807-802A-13 (1-736) x US-10-291-583-6 (1-4718)

Qy      1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20
Db      2223 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCAATTGCG 2282

Qy      21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp 40
Db      2283 GAGTGGTGGGACTTGAAACCTGGAGCCCGGAAGCCCAAGCCAAACCAAGCAAGCAGGAC 2342

Qy      41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp 60
Db      2343 GACGGCCGGGCTGGTGTCTTCTGGCTACAAGTACCTCGGACCTTCAACGGACTCGAC 2402

Qy      61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80
Db      2403 AAGGGGAGCCCGTCAACGGCGGACGAGCGGCCCTCGAGCAGCAAGGCCCTACGAC 2462

Qy      81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100
Db      2463 CAGCAGCTCAAGCGGGTGACAAATCCGTACCTGCGGTATAACACGCGCAGCGCCGAGTTT 2522

Qy      101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120
Db      2523 CAGGAGCGTCTGCAAGAAGATACGTCTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG 2582

Qy      121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGluGlyAlaLysThrAlaPro 140
Db      2583 GCCAAGAAGCGGGTTCTCGAACCTCTCGTCTGGTTGAGGAAGCGGCTAAGACGGCTCCT 2642

Qy      141 GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly 160
Db      2643 GGAAGAAGAACGTCGGTAGAGCAGTCGCCACAGAGCCAGACTCTCTCTCGGGCATCGGC 2702

Qy      161 LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180
Db      2703 AAGACAGGCCAGCAGCCCGCTAAAGAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG 2762

Qy      181 SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro 200
Db      2763 TCAGTCCCGATCCACAACCTCTCGGAGAACCTCCAGCAACCCCGCTGTGTGGGACCT 2822

Qy      201 ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly 220
Db      2823 ACTACAATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCCGACCGGA 2882

Qy      221 ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle 240
Db      2883 GTGGGTAATGCCTCAGGAAATTTGGCATTTGGCATTTCCACATGGCTGGGCGACAGAGTCATC 2942

Qy      241 ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle 260
Db      2943 ACCACCAGCACCCGACCTTGGCCCTTGGCCCTTACCAATACCAACCTCTACAAAGCAAATC 3002

Qy      261 SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp 280
Db      3003 TCCAGTGTCTTCAACGGGGGCCAGCAACGACACCACTACTTCCGCTACAGCACCCCTGG 3062

Qy      281 GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu 300
Db      4143 AAGAACCCCGCCTCCTCAGATCTCTCATCAAAACACGCGCTGTCTCTCGGAATCCTCCGGCG 4202
```

```
Db      3063 GGGTATTTTGATTTCACAGATTCCACTGCCACTTTTCCACCACGTGACTGCGCAGCGACTC 3122

Qy      301 IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln 320
Db      3123 ATCAACAACAATTTGGGGATTTCCGGCCCCAAGAGACTCAACTTCAAACCTTTCAACATCCAA 3182

Qy      321 ValLysGluValThrThrAsnAspGlyValThrIleAlaAsnAsnLeuThrSerThr 340
Db      3183 GTCAAGGAGGTACGACGAATGATGGCGTCACAAACCATCGCTAATAAACCTTTACCAGCAGC 3242

Qy      341 ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln 360
Db      3243 GTTCAAGTCTTCTCGGACTCGGAGTACCAAGTTCGCTACGTTCCGTCTCGGCTCTGCGCACCAG 3302

Qy      361 GlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr 380
Db      3303 GGCTGCCCTCCCTCCGTTCCCGCGGACGCTGTTTCATGATTCGCAATACGGCTACCTGACG 3362

Qy      381 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 400
Db      3363 CTCAACAATGGCAGCAAGCCGTTGGGACGTTTCATCTTTTACTGCTGGAATATTTCCCT 3422

Qy      401 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro 420
Db      3423 TCTCAGATGCTGAGAACGGGCAACAACCTTTACCTTCAGCTACACCTTTGAGGAAGTGCCT 3482

Qy      421 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 440
Db      3483 TTCCACAGCAGCTACGCGCACAGCCAGAGCCTGACCGGCTGATGAATCCTCTCATCGAC 3542

Qy      441 GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAsp 460
Db      3543 CAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAAGGAC 3602

Qy      461 LeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuPro 480
Db      3603 TTGCTGTTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTCAGCCCCAAAACACTGGCTACCT 3662

Qy      481 GlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsn 500
Db      3663 GGACCTGTATTATCGGCAGCAGCGCGTTTCTAAACAAAAAACAAGCAACAACAACAGCAAT 3722

Qy      501 PheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnPro 520
Db      3723 TTTACCTGGACTGGTGTCTCAAAATATAACCTCAATGGGCGTGAATCCCATCATCAACCTT 3782

Qy      521 GlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyVal 540
Db      3783 GGCACCTGCTATGGCCTCACACAAAGACGACGAAAGCAAGTTCTTTCCCATGAGCGGTGTC 3842

Qy      541 MetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIle 560
Db      3843 ATGATTTTGGAAAAGAGAGCGCCGGAGCTTCAAAACACTGCAATTTGGACAATGTTCATGATT 3902

Qy      561 ThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrVal 580
Db      3903 ACAGACGAAGAGGAAATTAAGCCACTAACCCCTGTGGCCACCCGAAAGATTTGGGACCGTG 3962

Qy      581 AlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGly 600
Db      3963 GCAGTCAATTTCCAGAGCAGCAGCAGACACGACCCCTGCGACCGGAGATGTGCATGTATGGGA 4022

Qy      601 AlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAla 620
Db      4023 GCATTACCTGGCATGGTGTGGCAAGATAGAGACGCTGTACCTGCAGGTTCCCATTTGGGCC 4082

Qy      621 LysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeu 640
Db      4083 AAAATTCTCTCACACAGATGGACACTTTTCAACCCGCTCTCTCTTATGGGCGGCTTTGGACTC 4142

Qy      641 LysAsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAla 660
Db      4143 AAGAACCCCGCCTCCTCAGATCTCTCATCAAAACACGCGCTGTCTCTCGGAATCCTCCGGCG 4202
```



QY 661 GluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSer 680  
Db 4203 GAGTTTTCAGCTACAAAGTTTGCTTCATTCATCAACCAATACTCCACAGGACAAGTGAGT 4262  
QY 681 ValGluIleGluTyrGluLeuGlnLysGluAsnSerLysArgTyrAsnProGluValGln 700  
Db 4263 GTGGAAATTGAATGGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAATCCGAAAGTGCAG 4322  
QY 701 TyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeu 720  
Db 4323 TACACATCCAATTATGCAAAATCTGCCAACGGTTGATTTTACTGTGGACAACAATGGACTT 4382  
QY 721 TyrThrGluProArgProIleGlyThrArgTyrIleuThrArgProLeu 736  
Db 4383 TATACTGAGCCTCGCCCATTTGGCACCCCGTTACCTTACCCGTCCCGCTG 4430  
RESULT 2  
US-10-216-870-11  
; Sequence 11, Application US/10216870  
; Publication No. US20030148506A1  
; GENERAL INFORMATION:  
; APPLICANT: KOTIN, ROBERT M  
; APPLICANT: URABE, MASASHI  
; APPLICANT: DING, CHUAN-TIAN  
; TITLE OF INVENTION: PRODUCTION OF ADENO-ASSOCIATED VIRUS IN INSECT CELLS  
; FILE REFERENCE: 402133  
; CURRENT APPLICATION NUMBER: US/10/216,870  
; CURRENT FILING DATE: 2002-08-13  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 11  
; LENGTH: 7447  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
US-10-216-870-11  
Alignment Scores:  
Pred. No.: 0 Length: 7447  
Score: 3984.00 Matches: 735  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.87% Indels: 0  
DB: 13 Gaps: 0  
US-09-807-802A-13 (1-736) x US-10-216-870-11 (1-7447)  
QY 2 AlaAlaAspGlyTyrLeuProAspTyrLeuGluAspAsnLeuSerGluGlyIleArgGlu 21  
Db 4624 GCTGCCGACGGTTATCTACCCGATTGGCTCGAGGACAACCTCTCTGAGGGCATTCGCGAG 4683  
QY 22 TrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAspAsp 41  
Db 4684 TGGTGGGACTTGAAACCTGGAGCCCCGAGCCCCAAGCCCAACCAGCAAAAGCAGGACGAC 4743  
QY 42 GlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAspLys 61  
Db 4744 GGCCGGGGTCTGGTGCTTCTGGCTACAAGTACCTCGGACCCCTTCAACGGACTCGACAAG 4803  
QY 62 GlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAspGln 81  
Db 4804 GGGGAGCCCCGTCAACCGCGGAGCGCAGCGGCCCTCGAGCACGACAAGGCCCTACGACCAG 4863  
QY 82 GlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPheGln 101  
Db 4864 CAGCTCAAAGCGGGTGACAATCCGTACCTGCGGTATAAACCAACGCGCAGCGCGAGTTTCAG 4923  
QY 102 GluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGlnAla 121  
Db 4924 GAGCGTCTGCAAGAAGATACGTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAGGCC 4983

QY 122 LysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaProGly 141  
Db 4984 AAGAAGCGGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGCTCCTTGA 5043  
QY 142 LysLysArgProValGluGlnSerProGlnGluProAspSerSerGlyIleGlyLys 161  
Db 5044 AAGAAACGTCCGGTAGAGCAGTCCCAAGAGCCAGACTCCTCTCGGCATCGGCAAG 5103  
QY 162 ThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGluSer 181  
Db 5104 ACAGGCCAGCAGCCCGCTAAAGAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAGTCA 5163  
QY 182 ValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyProThr 201  
Db 5164 GTCCCCGATCCACAACCTCTCGGAGAACCTCCAGCAACCCCGCTGCTGTGGACCTACT 5223  
QY 202 ThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyVal 221  
Db 5224 ACAATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGCGCCGCGGAGTG 5283  
QY 222 GlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThr 241  
Db 5284 GGTAATGGCTCAGGAATTTGGCAATTGGCAATTCGATTCCACATGGCTGGCGACAGAGTCATCACC 5343  
QY 242 ThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSer 261  
Db 5344 ACCAGCACCCGACCTGGGCTTGGCCACCTACAATAACCACTCTACAGCAAAATCTCC 5403  
QY 262 SerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTyrGly 281  
Db 5404 AGTGCTTCAACGGGGGCCAGCAACGACAACCACTACTTCGGCTACAGCACCCCTGGGGG 5463  
QY 282 TyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIle 301  
Db 5464 TATTTTGATTTCAACAGATTCCACTGCCACTTTTCCACCAGTGACTGGCAGCGACTCATC 5523  
QY 302 AsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnVal 321  
Db 5524 AACAAACAATTGGGGATTCCGGGCCCAAGAGACTCAACTTCAAACTTTCAACATCCAAAGTC 5583  
QY 322 LysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrVal 341  
Db 5584 AAGGAGGTACGACGAGAAATGATGGCGTCACAACCATCGCTAATAACCTTACCAGCACGGTT 5643  
QY 342 GlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGly 361  
Db 5644 CAAGTCTTCTCGGACTCGGAGTACCAGCTTCCGTACGTCTCGCTCGGCTGTGCGCACCAGGGC 5703  
QY 362 CysLeuProPropheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeu 381  
Db 5704 TGCTCTCCCTCCGTTCCCGCGGACGTGTTCATGATTCGCAATACGGCTACCTGACGCTC 5763  
QY 382 AsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSer 401  
Db 5764 AACAAATGGCAGCCCAAGCCGTGGGACGTTTCATCCTTTTACTGCTGGAATATTTCCCTTCT 5823  
QY 402 GlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProphe 421  
Db 5824 CAGATGCTGAGAACGGGCAACAACCTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTC 5883  
QY 422 HisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGln 441  
Db 5884 CACAGCAGCTACGCGCACAGCCAGAGCTGGACCGGTGATGATGATCTCTCATCGACCAA 5943  
QY 442 TyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeu 461  
Db 5944 TACCTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCTGGAAGCAAGGACTTG 6003  
QY 462 LeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGly 481  
Db 6004 CTGTTTAGCCGTGGGTCTCCAGCTGSCATGTCTGTTCAGCCCCCAAAACTGGCTACCTTGA 6063  
QY 482 ProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPhe 501

```
Db 6064 CCCTGTATCGGCAGCGCGTTTCTAAACAAACAGCAAAACAGCAATTTT 6123
QY 502 ThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGly 521
Db 6124 ACCTGGACTGGTGCTTCAAAATATAAACCTCAATGGCGTGAATCCATCATCAACCTGGC 6183
QY 522 ThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMet 541
Db 6184 ACTGCTATGGCCTCACACAAAGACGACGAAGATTCTTCCCATGAGCGGTGTCATG 6243
QY 542 IlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThr 561
Db 6244 ATTTTGGAAAGAGAGCGCGGAGCTTCAACACTGCATGTGACAATGTTCATGATTACA 6303
QY 562 AspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAla 581
Db 6304 GACGAAGAGAAATTAAAGCCACTTAACCCTGTGGCCACCGAAGATTTTGGGACCGTGGCA 6363
QY 582 ValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAla 601
Db 6364 GTCAATTTCCAGAGCAGCAGCACAGACCCTGCGACCGGAGATGTGCATGCTATGGGAGCA 6423
QY 602 LeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLys 621
Db 6424 TTACCTGGCATGGTGTGGCAAGATAGACGCTGTACCTGCAGGCTCCCATTTGGGCCAAA 6483
QY 622 IleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLys 641
Db 6484 ATTCTCTCACAGATGGACACTTTTACCCGCTCTCTCTTATGGCGGCTTTGGACTCAAG 6543
QY 642 AsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGlu 661
Db 6544 AACCCGCCCTCTCAGATCCTCATCAAAACACAGCCTGTTCTCTGCGAATCCTCCGCGGAG 6603
QY 662 PheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerVal 681
Db 6604 TTTTTCAGTACAAAGTTTGTTCATTTCATCCCAATACTCCACAGGACAAAGTGAGTGTG 6663
QY 682 GluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyr 701
Db 6664 GAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTAC 6723
QY 702 ThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyr 721
Db 6724 ACATCCAAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAACAATGGACTTTAT 6783
QY 722 ThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736
Db 6784 ACTGAGCCTCGCCCATTTGGCACCCGTTACCTTACCCGTCCCTG 6828

RESULT 3
US-10-291-583-8
; Sequence 8, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1 Sequences Identifi
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 4726
; TYPE: DNA
; ORGANISM: adeno-associated virus serotype 3
US-10-291-583-8

Alignment Scores:
Pred. No.: 0 Length: 4726
Score: 3494.00 Matches: 637
Percent Similarity: 91.72% Conservative: 39
Best Local Similarity: 86.43% Mismatches: 59
Query Match: 87.59% Indels: 2
DB: 13 Gaps: 2

US-09-807-802A-13 (1-736) x US-10-291-583-8 (1-4726)

QY 1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20
Db 2209 ATGGCTGCTGACGGTTATCTTCCAGATTGGCTCGAGGACAACCTTTTCTGAAGGCATTCGT 2268
QY 21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp 40
Db 2269 GAGTGTGGGCTCTGAACCTGGAGTCCCTCAACCCAAAGCGAACCAACACACCCAGGAC 2328
QY 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp 60
Db 2329 AACCGTCGGGCTCTTGTGCTTCCGGGTTACAAATACCTCGGACCCGGTAAACGAGCTCGAC 2388
QY 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80
Db 2389 AAAGGAGAGCCGCTCAACGAGGCGGACGCGGAGCCCTCGAACACGACAAAGCTTACGAC 2448
QY 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100
Db 2449 CAGCAGCTCAAGCCGCTGACAAACCGGTACCTCAAGTACAAACACGCGGACGCGAGTTT 2508
QY 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120
Db 2509 CAGGAGCGTCTTCAAGAAGATACGTCTTTTGGGGCAACCTTGGCAGAGCAGTCTTCCAG 2568
QY 121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGluGlyAlaLysThrAlaPro 140
Db 2569 GCCAAAAGAGGATCCTTGAGCCTCTTGGTCTGGTTGAGGAAGCAGCTAAACCGGTCCT 2628
QY 141 GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly 160
Db 2629 GGAAAGAAAGGGGCTGTAGATCAGTCTCCTCAGGAACCGGACTCATCTCTGGTGTGGC 2688
QY 161 LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180
Db 2689 AAATCGGGCAACAGCCTGCCAGAAAAGACTAAATTTCCGTTCAGACTGGAGACTCAGAG 2748
QY 181 SerValProAspProGlnProLeuGluProProAlaThrProAlaAlaValGlyPro 200
Db 2749 TCAGTCCCAGACCCCTCAACCTCTCGGAGAACCCACCAGCAGCCCCCAACAGTTTGGGATCT 2808
QY 201 ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly 220
Db 2809 AATACAATGGCTTCAGCGCGTGGCGCACCAATGGCAGACAATAACGAGGGTGCCGATGGA 2868
QY 221 ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle 240
Db 2869 GTGGGTAATTCCTCAGGAATTTGGCATTTCCGATTTCCCAATGGCTGGCGGACAGAGTCATC 2928
QY 241 ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle 260
Db 2929 ACCACGACGACCAGAACCTGGGCCCTGCCCACTTACAAACAACCATCTCTACAAGCAAAATC 2988
QY 261 SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp 280
Db 2989 TCCAGCCAATCA---GGAGCTTCAAAACGACAACCACTACTTTGGCTACAGCACCCCTTGG 3045
```

Qy 281 GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu 300  
Db 3046 GGGTATTTTGACTTTAACAGATTCCACTGCCACTTCTCACCACGTGACTGGCAGGACTC 3105  
Qy 301 IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln 320  
Db 3106 ATTAAACAACAACCTGGGGATTCCGGCCCAAGAAACTCAGCTTCAAGCTCTTCAACATCCAA 3165  
Qy 321 ValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThr 340  
Db 3166 GTTAGAGGGGTCAACGAGATGGCAGCAGCTATTCGCCAATAACCTTTACCAGCAGC 3225  
Qy 341 ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln 360  
Db 3226 GTTCAAGTGTTTACGGACTCGGAGTATCAGCTCCCGTAGCTGCTCGGGTCGGCGCACCA 3285  
Qy 361 GlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr 380  
Db 3286 GGCTGTCTCCCGCCGTTTCCAGCGGACGTCTTCATGGTCCCTCAGTATGGATACCTCACC 3345  
Qy 381 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 400  
Db 3346 CTGAACAACGGAAAGTCAAGCGGTGGACGCTCATCTTTTACTGCTGGAGTACTTCCCT 3405  
Qy 401 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro 420  
Db 3406 TCGCAGATGCTAAGGACTGGAAATAACTTCCAATTTCAGCTATACCTTCGAGGATGACCT 3465  
Qy 421 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 440  
Db 3466 TTTCACAGCAGCTACGCTCACAGCAGAGTTTGGATCGCTTGATGAATCCTCTTATTGAT 3525  
Qy 441 GlnTyrLeuTyrTyrLeuAsnArgThrGln---AsnGlnSerGlySerAlaGlnAsnLys 459  
Db 3526 CAGTATCTGTACTACCTGAACAGAACGCAAGGAACAACCTCTGGAACAACAACCAATCA 3585  
Qy 460 AspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeu 479  
Db 3586 CGGCTGCTTTTAGCCAGGCTGGGCTCAGTCTATGTCTTTGAGGCCAGAAATTGGCTA 3645  
Qy 480 ProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSer 499  
Db 3646 CCTGGGCCCTGTACCGGGCAACAGAGACTTTCAAAGACTGCTAACGACAACAACACAGT 3705  
Qy 500 AsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsn 519  
Db 3706 AACTTTCCTTGGACAGCGGCCAGCAAAATATCATCTCAATGGCCGCGACTCGCTGGTGAAT 3765  
Qy 520 ProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGly 539  
Db 3766 CCAGGACCAGCTATGGCCAGTCACAAGGACGATGAAGAAAAATTTTCCCTATGCACGGC 3825  
Qy 540 ValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMet 559  
Db 3826 AATCTAATATTGGCAAGAAGGGACAACGGCAAGTAACGACAGAAATTAGATAATGTAATG 3885  
Qy 560 IleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThr 579  
Db 3886 ATTACGGATGAAGAAGAGATTTCGTACCACCAATCCTGTGGCAACAGACGATATGGAAC 3945  
Qy 580 ValAlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMet 599  
Db 3946 GTGGCAAAATAACTTCAGAGCTCAAATACAGCTCCCACGACTGGAACGTGTCATCATCAG 4005  
Qy 600 GlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrp 619  
Db 4006 GGGGCCTTACCTGGCATGGTGTGGCAAGATCGTGACGTGTACCTTCAAGGACCTACTCTGG 4065  
Qy 620 AlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGly 639  
Db 4066 GCAAAGATTCTCACACGGATGGACACTTTCATCTCTCTCTCTGATGGGAGGCTTTGGA 4125  
Qy 640 LeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProPro 659

Db 4126 CTGAACATCCGCTCCTCAATCATGATCAAAAAATACTCCGGTACCGGCAAAATCCTCCG 4185  
Qy 660 AlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnVal 679  
Db 4186 ACGACTTTTCAGCCCGGCCAAGTTTGCTTCATTATTCACCTCAGTACTCCACTGCACAGGTC 4245  
Qy 680 SerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTyrAsnProGluVal 699  
Db 4246 AGCGTGGAAATTGAGTGGGAGCTACAGAAAAGAAAACAGCAACGTTGGGAATCCAGAGATT 4305  
Qy 700 GlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGly 719  
Db 4306 CAGTACACTTCCAACTACAACAAGTCTGTTAATGTGGACTTTTACTGTAGACACTAATGGT 4365  
Qy 720 LeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736  
Db 4366 GTTATAGTGAACCTCGCCCTATTGGAACCCCGTATCTCACACGAAACTTG 4416  
RESULT 4  
US-10-291-583-45  
; Sequence 45, Application US/10291583  
; Publication No. US20030138772A1  
; GENERAL INFORMATION:  
; APPLICANT: Gao, Guangping  
; APPLICANT: Wilson, James M.  
; APPLICANT: Alvira, Mauricio  
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus  
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1  
; FILE REFERENCE: UPN-02735USA  
; CURRENT APPLICATION NUMBER: US/10/291,583  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: US 60/350,607  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/341,117  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: US 60/377,066  
; PRIOR FILING DATE: 2002-05-01  
; PRIOR APPLICATION NUMBER: US 60/386,675  
; PRIOR FILING DATE: 2002-06-05  
; NUMBER OF SEQ ID NOS: 120  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 45  
; LENGTH: 3122  
; TYPE: DNA  
; ORGANISM: 43.25  
US-10-291-583-45  
Alignment Scores:  
Pred. No.: 0 Length: 3122  
Score: 3473.00 Matches: 631  
Percent Similarity: 91.59% Conservative: 44  
Best Local Similarity: 85.62% Mismatches: 60  
Query Match: 87.06% Indels: 2  
DB: 13 Gaps: 2  
US-09-807-802A-13 (1-736) x US-10-291-583-45 (1-3122)  
Qy 1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20  
Db 844 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTCGC 903  
Qy 21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp 40  
Db 904 GAGTGGTGGGACTTGAAACCTGGAGCCCCGAAACCCAAAGCCAAACAGCAAAAGCAGGAC 963  
Qy 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp 60  
Db 964 GACGGCCGGGGTCTGGTGTCTTCTGGCTACAAGTACCTTCGGACCCCTTCACGCGACTCGAC 1023  
Qy 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80  
Db 1024 AAGGGGAGCCCCGTCAACGCGGCGACGCGCGGCCCTTCGAGCACGACAAAGCCTACGAC 1083



QY 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100  
Db 1084 CAGCAGCTCAAAGCGGGTGACAAATCCGTACCTGCGGTATAATACACGCCGACGCCGAGTTT 1143  
QY 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120  
Db 1144 CAGGAGCGTCTGCAAGAAGATACGTCCTTTGGGGGCAACCTCGGGCGAGCAGTCCTCCAG 1203  
QY 121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro 140  
Db 1204 GCCAAGAGCGGGTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGGCTAAGACGGCTCCT 1263  
QY 141 GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly 160  
Db 1264 GGAAAGAGAGACCGGTAGAGCAGTCGCCACAAGAGCCAGACTCCTCCTCGGGCATCGGC 1323  
QY 161 LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180  
Db 1324 AAGACAGCGCAGCAGCCCGCTAAAGAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG 1383  
QY 181 SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro 200  
Db 1384 TCAGTCCCGACCCACAACCTCTCGGAGAACCTCCAGCAGCCCGCTCAGGTCTGGGACCT 1443  
QY 201 ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly 220  
Db 1444 AATACAATGGCTTCAGGCGGTGGCGCTCCAATGGCAGACAAATAACGAAGCGCGCACGGA 1503  
QY 221 ValGlyAsnAlaSerGlyAsnTyrHisCysAspSerThrTyrLeuGlyAspArgValIle 240  
Db 1504 GTGGGTAAITCTCGGGAATTTGGCAATTTGGCATTTCCACATGGCTGGGGACAGAGTCATC 1563  
QY 241 ThrThrSerThrArgThrTyrAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle 260  
Db 1564 ACCACCAGCACCGAACCTGGGCCCTGCCACCTACAACAACCACTCTACAAGCAATC 1623  
QY 261 SerSerAlaSerThrGlyAlaSer---AsnAspAsnHisTyrPheGlyTyrSerThrPro 279  
Db 1624 TCCAACGGCCTCGGGAGGAAGCACCAACGACCAACACCTATTTTGGCTACAGCACCCCC 1683  
QY 280 TrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArg 299  
Db 1684 TGGGGGTATTTTGACTTCAACAGATTCCACTGTCACTTTTCCACCAGTGTGGCAACGA 1743  
QY 300 LeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIle 319  
Db 1744 CTCAATCAACAACAATTTGGGATTTCCGGCCCAAAAGACTCAACTTCAAGCTGTTCAACATC 1803  
QY 320 GlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSer 339  
Db 1804 CAGGTCAAGGAAGTCACGACGACGAACGAAGGCCAACCAAGACCATCGCCAATAATCTCACCAGC 1863  
QY 340 ThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHis 359  
Db 1864 ACCGTGCAGGTCTTTACGGACTCGGAGTACCAGTTACCGTACGTGCTAGGATCCGCTCAC 1923  
QY 360 GlnGlyCysLeuProPropheProAlaAspValPheMetIleProGlnTyrGlyTyrLeu 379  
Db 1924 CAGGGATGTCTGCTCCGTTCCCGCGGAGCTTTCATGGTTCTCCTCAGTACGGCTATTTA 1983  
QY 380 ThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhe 399  
Db 1984 ACTTTAAACAATGGAAGCCCAAGCCCTGGGACGTTCTCCTCTACTGTCTGGAGTATTTT 2043  
QY 400 ProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluVal 419  
Db 2044 CCATCGCAGATGCTGAGAACCGGCAACAACCTTCAGTTTCAGTTACGCTACGCTTCGAGGACGTG 2103  
QY 420 PropheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIle 439  
Db 2104 CCTTTCCACAGCAGCTACCGGCACAGCCAGAGCTGGACAGGCTGATGAATCCCTCATC 2163

QY 440 AspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLys 459  
Db 2164 GACCAGTACCTGTACTACCTGGTCAGAACGCAACG--ACTGGAACCTGGAGGACGCGAG 2220  
QY 460 AspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeu 479  
Db 2221 ACTCTGGCATTCAGCAAGCGGGTCTTAGCTCAATGGCCAACCGAGGTAGAAATTTGGGTG 2280  
QY 480 ProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnSer 499  
Db 2281 CCGGACCTTGTACCGGCAGCAGCGCGTCTCCAGCACAACCAACCAACCAACACAGC 2340  
QY 500 AsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsn 519  
Db 2341 AACTTTGCCTGGACGGGAGCTGCCAAGTTTAAAGCTGAACGGCCGAGACTCTCTAATGAAT 2400  
QY 520 ProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGly 539  
Db 2401 CCGGGCGTGGCAATGGCTTCCACAAGGATGACGACGACCGCTTCTTCCCTTCGAGCGGG 2460  
QY 540 ValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMet 559  
Db 2461 GTCTGTGATTTTGGCAAGCAAGGAGCGCGGAACGATGGAGTGGATTACAGCCAAAGTGTCTG 2520  
QY 560 IleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThr 579  
Db 2521 ATTACAGATGAGGAAGAATCAAGGCTACCAACCCCGTGGCCACAGAGAATATATGGAGCA 2580  
QY 580 ValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMet 599  
Db 2581 GTGGCCATCAACAACCGAGCGCCCAATACGCGCGCAGACCGGACTCGTGCACAACCCAG 2640  
QY 600 GlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrp 619  
Db 2641 GGGGTGATTTCCCGCATGGTGTGGCAGAATAGACAGTGTACCTGCAGGGTCCCATCTGG 2700  
QY 620 AlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGly 639  
Db 2701 GCCAAAATTCCTCACCGGACGGCAACTTTACCCGCTCTCCCTGTATGGCGCGCTTTGGA 2760  
QY 640 LeuLysAsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProPro 659  
Db 2761 CTGAAGCACCCCGCTCTCTCAAAATTTCTCATCAAGAAACACACCGGTTCCAGCGGACCCGCG 2820  
QY 660 AlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnVal 679  
Db 2821 CTTACCTTCAACCAGGCCCAAGCTGAACCTCTTTCATCAGCAGTACAGCACCGGACAGGTC 2880  
QY 680 SerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluVal 699  
Db 2881 AGCGTGGAAATCGAGTGGAGCTGCAGAAAGAAACAGCAAAACGCTGGAATCCAGAGATT 2940  
QY 700 GlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGly 719  
Db 2941 CAATACACTTCCAACCTACTACAAATCTACAAATGTGGACTTTGCTGTCAACACCGSAGGG 3000  
QY 720 LeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736  
Db 3001 GTTATAGCGAGCCTCGCCCATTTGGCACCCCGTTACTCTACCCCGCAACCTG 3051

RESULT 5

US-10-291-583-43  
; Sequence 43, Application US/10291583  
; Publication No. US20030138772A1  
; GENERAL INFORMATION:  
; APPLICANT: Gao, Guangping  
; APPLICANT: Wilson, James M.  
; APPLICANT: Alvira, Mauricio  
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus  
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1e1 Sequences Identif  
; FILE REFERENCE: UPN-02735USA  
; CURRENT APPLICATION NUMBER: US/10/291,583  
; CURRENT FILING DATE: 2002-11-12

; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43
; LENGTH: 3117
; TYPE: DNA
; ORGANISM: 43.21
US-10-291-583-43

Alignment Scores:

Pred. No.: 0 Length: 3117
Score: 3464.00 Matches: 629
Percent Similarity: 91.59% Conservative: 46
Best Local Similarity: 85.35% Mismatches: 60
Query Match: 86.84% Indels: 2
DB: 13 Gaps: 2

US-09-807-802A-13 (1-736) x US-10-291-583-43 (1-3117)

QY 1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20
DB 839 ATGGCTGCCGATGGTTATCTCCAGATTGGCTCGAGGACAACCTCTCTGAGGCGATTCCG 898
QY 21 GluTrpTyrAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp 40
DB 899 GAGTGGTGGGACTTGAACCTGGAGCCCCGAAACCCAAAGCCAAACCCAGCAAGCAGGAC 958
QY 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp 60
DB 959 GACGGCCGGGTCTGGTGTCTCTGGCTACAAAGTACCTCGGACCCCTTCAACGGGACTCGAC 1018
QY 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80
DB 1019 AAGGGGAGCCCGTCAACGGCGGAGCGACGGCCCTCGAGCACGACAAAGCCTACGAC 1078
QY 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100
DB 1079 CAGCAGCTCAAAGCGGGTGACAATCCGTACCTGCGGTATTAATCACGCCGACGCCGAGTTT 1138
QY 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120
DB 1139 CAGGAGCGTCTGCAAGAGATACGTCTTTGGGGGCACTCTGGGCGAGCAGTCTTCCAG 1198
QY 121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro 140
DB 1199 GCCAAGAAGCGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGTCCT 1258
QY 141 GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly 160
DB 1259 GGAAAGAAGAGACCGGTAGAGCAGTGCACCAAGAGCCAGACTCCTCTCGGGCATCGGC 1318
QY 161 LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180
DB 1319 AAGACAGGCCAGCAGCCCGCTTAAAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG 1378
QY 181 SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro 200
DB 1379 TCAGTCCCCGACCCACACCTCTCGGAGAACCTCCAGCAGCCCCCTCAGGTCTGGGACCT 1438
QY 201 ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly 220
DB 1439 AATACAATGGCTTCAGGCGGTGGCGCTCCAATGGCAGACATAACGAAGCGCCGACGGA 1498
QY 221 ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle 240
DB 1499 GTGGGTAAATCTCTCGGGAAATTTGGCATTCGATTCCACATGGCTGGGGGACAGAGTCATC 1558

QY 241 ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle 260
DB 1559 ACCACCAGCACCCGAACCTGGGCCCTGCCCCACCTACAACAACCACTCTACAAAGCAAATC 1618
QY 261 SerSerAlaSerThrGlyAlaSer---AsnAspAsnHisTyrPheGlyTyrSerThrPro 279
DB 1619 TCCAACGGCACCTCGGAGGAAGCAACCAACGACAACACCTATTTTGGGTACAGCACCCCC 1678
QY 280 TrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArg 299
DB 1679 TGGGGGTATTTGACTTCAACAGATCCACTGTCACTTTTCAACCACGTGACTGGCAACGA 1738
QY 300 LeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIle 319
DB 1739 CTCATCAACAACAATTTGGGATTCCGGCCCAAAAGACTCAACTTCAAGCTGTTCAACATC 1798
QY 320 GlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSer 339
DB 1799 CAGGTCAAGGAAGTCACGACGAACGAAGGACCAAGACCATCGCCAATAATCTCACCAGC 1858
QY 340 ThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHis 359
DB 1859 ACCGTGCGGTCTTTACGGACTCGGAGTACCAAGTTACCGTACGTGTAGGATCCGCTCAC 1918
QY 360 GlnGlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeu 379
DB 1919 CAGGGATGCTTCCTCCGTTCCCGGCGGACGCTTTCATGGTTCCTCAGTACGGCTATTTA 1978
QY 380 ThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhe 399
DB 1979 ACTTTAAACAATGGAAGCCCAAGCCCTGGGACGTTCTCTCTACTGTCTGGAGTATTTT 2038
QY 400 ProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluVal 419
DB 2039 CCATCGCAGATGCTGAGAACCGGCAACAACATTTTCAGTTTCAGTACACCTTCGAGGACGTG 2098
QY 420 PropheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIle 439
DB 2099 CCTTTCCACAGCAGCTACGCGCACAGCCAGAGCCTGGACAGGCTGATGAATCCCTCATC 2158
QY 440 AspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLys 459
DB 2159 GACCAGTACCTGTACTACCTGGTCAGAACGCAAAACG--ACTGGAACTGGAGGAGCGCAG 2215
QY 460 AspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeu 479
DB 2216 ACTCTGGCATTCAGCCAAGCGGGTCTTAGCTCAATGGCCAACCAAGGCTAGAAATTTGGGTG 2275
QY 480 ProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSer 499
DB 2276 CCCGGACCTTGCTACCGGCAGCAGCGCGTCTCCAGACAACCAACCAAGCAACACAGC 2335
QY 500 AsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsn 519
DB 2336 AACTTTGCTGGAGCGGAGCTGCCAAGTTTAAAGCTGAACGGCCGAGACTCTCTAATGAAT 2395
QY 520 ProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGly 539
DB 2396 CCGGGCGTGGCAATGGCTCCCAACAGGATGACGACGACCGCTTCTTCCCTTCGAGCGGG 2455
QY 540 ValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMet 559
DB 2456 GTCTGTGATTTTGGCAAGCAAGGAGCCGGGACGATGGAGTGGATTACAGCCAAAGTGTCTG 2515
QY 560 IleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThr 579
DB 2516 ATTACAGATGAGGAAGAAATCAAGGCTACCAACCCCGTGGCCACAGAGAATATGGAGCA 2575
QY 580 ValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMet 599
DB 2576 GTGGCCATCAACAACAGGCGCCCAATACGCGCGCAGACCGGACTCGTGCACAACACCAG 2635

```
Qy 600 GlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrp 619
    |||:::|||||
Db 2636 GGGTGATCCCGGCATGGTGTGCAGAAATAGAGACGTGTACTGCAGGGTCCCATCTGG 2695

Qy 620 AlaValIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGly 639
    |||:::|||||
Db 2696 GCCAAATTCCTCACACGGACGGCAACTTTCACCCGTCTCCCTGATGGCGGCTTGG 2755

Qy 640 LeuLysAsnProProProGlnIleLeuLysAsnThrProValProAlaAsnProPro 659
    |||:::|||||
Db 2756 CTGAAGCACCCCGCTCTCTCAAAATTCATCAAGAACACACCGGTTCAGCGGACCCGCG 2815

Qy 660 AlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnVal 679
    |||:::|||||
Db 2816 CTTACCTTCAACAGGCAAGCTGAACCTTTTCATCACGCGAGTACAGCACCGGACAGTC 2875

Qy 680 SerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluVal 699
    |||:::|||||
Db 2876 AGCGTGGAAATCGAGTGGGAGCTGCAGAAAGAAACAGCAACCGCTGGAATCCAGAGATT 2935

Qy 700 GlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGly 719
    |||:::|||||
Db 2936 CAATACACTTCCAACACTACTACAAATCTACAAATGTGGACTTTGCTGTCAACACGGGAAG 2995

Qy 720 LeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736
    |||:::|||||
Db 2996 GTTTATAGCGAGCCTCGCCCCATTGGCACCCCTTACCTCACCCGCAACCTG 3046

RESULT 6
US-10-291-583-44
; Sequence 44, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; PRIORITY FILING DATE: 2002-11-12
; PRIORITY FILING DATE: 2002-11-12
; PRIORITY FILING DATE: 2001-11-13
; PRIORITY FILING DATE: 2001-12-17
; PRIORITY FILING DATE: 2001-12-17
; PRIORITY FILING DATE: 2002-05-01
; PRIORITY FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 3121
; TYPE: DNA
; ORGANISM: 43.23
US-10-291-583-44

Alignment Scores:
Pred. No.: 0 Length: 3121
Score: 3460.00 Matches: 629
Percent Similarity: 91.32% Conservative: 44
Best Local Similarity: 85.35% Mismatches: 62
Query Match: 86.74% Indels: 2
DB: 13 Gaps: 2

US-09-807-802A-13 (1-736) x US-10-291-583-44 (1-3121)

Qy 1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20
    |||:::|||||
Db 843 ATGGCTGCCGATGGTTATCTCCAGATTGCTCGAGGACAACCTCTCTGAGGGCATTCCG 902

Qy 21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnLysGlnAsp 40
    |||:::|||||
```

```
Db 903 GAGTGGTGGGACTTGAAACCTGGAGCCCCCGAAACCCCAAGCCAAACCCAGCAAAAGCAGGAC 962

Qy 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp 60
    |||:::|||||
Db 963 GACGGCCGGGGTCTGGTGCTTCTCTGGCTACAAGTACCTCGGACCCCTTCAACGCGACTCGAC 1022

Qy 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80
    |||:::|||||
Db 1023 AAGGGGAGCCCCGTCAACGCGCGGACGCGCGGCCCTCGAGCACGACAAAGCCTACGAC 1082

Qy 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100
    |||:::|||||
Db 1083 CAGCAGCTCAAAGCGGGTGACAAATCCGTACCTCGGTATATAATCACGCCGACGCCGAGTTT 1142

Qy 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120
    |||:::|||||
Db 1143 CAGGAGCGTCTGCAAGAAGATACGTCTCTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG 1202

Qy 121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro 140
    |||:::|||||
Db 1203 GCCAAGAAGCGGGTTCGAACCTCTCGGTCTGTGTGAGGAAGCGCTAAGACGGCTCCT 1262

Qy 141 GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly 160
    |||:::|||||
Db 1263 GGAAAGAAGAGACCCGTAGAGCAGTCGCCACACAGAGCCAGACTCTCTCTCGGGCATCGGC 1322

Qy 161 LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180
    |||:::|||||
Db 1323 AAGACAGGCCAGCAGCCCGCTAAAAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG 1382

Qy 181 SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro 200
    |||:::|||||
Db 1383 TCAGTCCCCGACCCCAACCTCTCGGAGAACCTTCAGAGCCCCCTCAGGCTTGGGACCT 1442

Qy 201 ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly 220
    |||:::|||||
Db 1443 AATACAATGGCTTCAGCGGTGGCGCTCCAAATGGCAGACAATAACGAAGCGCGCACGGA 1502

Qy 221 ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle 240
    |||:::|||||
Db 1503 GTGGGTAATTCCTCGGAAATTTGGCATTTGCGATTCCACATGGCTGGGGGACAGAGTCATC 1562

Qy 241 ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle 260
    |||:::|||||
Db 1563 ACCACCAGCACCCGAACTGGGCCCTGCCACCTACAAACAACCACTCTACAAAGCAAATC 1622

Qy 261 SerSerAlaSerThrGlyAlaSer--AsnAspAsnHisTyrPheGlyTyrSerThrPro 279
    |||:::|||||
Db 1623 TCCAACGGCACCTCGGAGGAAGCACCAACGACGACACACCTATTTTGGCTACAGCACCCCC 1682

Qy 280 TrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArg 299
    |||:::|||||
Db 1683 TGGGGGTATTTTGACTTCAACAGATTCCACTGTCTACTTTTCAACCACGCTGCTGGCAACGA 1742

Qy 300 LeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIle 319
    |||:::|||||
Db 1743 CTCATCAACAACAATTTGGGATTTCCGGCCCCAAAAGACTCAACTTCAAGCTGTTCAACATC 1802

Qy 320 GlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSer 339
    |||:::|||||
Db 1803 CAGGTCAAGGAAGTACAGCAGAAACGAAGGCACCAAGACCATCGCCAATAATCTCACGAC 1862

Qy 340 ThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHis 359
    |||:::|||||
Db 1863 ACCGTGCAGGTCTTTACGAGACTTGGAGTACCAGTTACGTTACGTTACGTTACGTTACGTTAC 1922

Qy 360 GlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeu 379
    |||:::|||||
Db 1923 CAGGGATGTCTGCCTCCGTTCGCGGGGACGCTCTTCATGGTTCTCTCAGTACGGCTATTTA 1982

Qy 380 ThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhe 399
    |||:::|||||
Db 1983 ACTTTAAACAATGGAAAGCCCAAGCCCTGGGACGTTCTCTCTCTACTGTCTGGAGTATTTC 2042
```





Db 1384 TCAGTCCCCGACCCACAACTCTCGGAGAACTCCAGCAGCCCCCTCAGGTCGGGACCT 1443  
Qy 201 ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly 220  
Db 1444 AATACAATGGCTTCAGCGGTGGCGCTCCAATGGCAGACAATAACGAAGCGCGACGGA 1503  
Qy 221 ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle 240  
Db 1504 GTGGGTAATTCCTCGGGAATTTGGCATTTGCATTTCCAGATTCCACATGGCTGGGGACAGAGTCATC 1563  
Qy 241 ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle 260  
Db 1564 ACCACGACACCCGAACCTGGGCCCTGCCACCTACAACAACCCACCTCTACAAGCAAAATC 1623  
Qy 261 SerSerAlaSerThrGlyAlaSer--AsnAspAsnHisTyrPheGlyTyrSerThrPro 279  
Db 1624 TCCACGGCACCTCGGAGGAGACCAACGACAACACCTATTTTGGCTACAGCACCCCC 1683  
Qy 280 TrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArg 299  
Db 1684 TGGGGGTATTTTGACTTCAACAGATTCCACTGTCCACTTTTCCACCACGTGACTGGCAACGA 1743  
Qy 300 LeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIle 319  
Db 1744 CTCATCAACAACAATTTGGGATTCCGGCCCCAAAGACTCAACTTCAAGCTGTTCAACATC 1803  
Qy 320 GlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSer 339  
Db 1804 CAGGTCAAGGAAGTCACGACGACGAAGAGGCACCAAGACCATTGCGCAATAATCTCACACGC 1863  
Qy 340 ThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHis 359  
Db 1864 ACCGTGCAGTCTTTACGGACTCGGAGTACCAGTTACCGTACGTGTAGGATCCGCTCAC 1923  
Qy 360 GlnGlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeu 379  
Db 1924 CAGGGATGTCCTCCGTTCCCGCGGACGTCTTCACGGTTCCTCAGTACGGCTATTTA 1983  
Qy 380 ThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhe 399  
Db 1984 ACTTTAAACAATGGAAGCCCAAGCCCTGGGACGTTCTCTCTCTACTGTCTGGAGTATTC 2043  
Qy 400 ProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluVal 419  
Db 2044 CCATCGCAGATGCTGAGAACCGGCAACAACTTTCAGTTTCAGTTCACCTTCGAGGACGTG 2103  
Qy 420 PropheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIle 439  
Db 2104 CCTTTCCACAGAGCTACGCGCAGCAGCCAGCGCTGGACAGCGCTGGAGGTGAATCCCTCATC 2163  
Qy 440 AspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLys 459  
Db 2164 GACCAGTACCTGTACTACCTGGTCAGAACGCAAAACG---ACTGGAACTGGAGGACGCAG 2220  
Qy 460 AspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeu 479  
Db 2221 ACTCTGGCATTCAGCAAGCGGGTCTTAGCTCAATGGCCAACCAAGGCTAGAAATTGGGTG 2280  
Qy 480 ProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnSer 499  
Db 2281 CCCGGACCTTGCTACCGGCAGCAGCGCGTCTCCAGCACAACCAACCAACAACACAGC 2340  
Qy 500 AsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsn 519  
Db 2341 AACTTTGCCCTGGACGGGAGCTGCCAAGTTTAAGCTGAACGGCCGAGACTCTCTAATGAAT 2400  
Qy 520 ProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGly 539  
Db 2401 CCGGGCGTGGCAATGGCTTCCACAAGGATGACGACGACCGCTTCTTCCCTTCGAGCGGG 2460  
Qy 540 ValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMet 559  
Db 2461 GTCTTGATTTTGGCAAGCAAGGAGCGCGGAAACGATGGAGTGGATTACAGCCAAGTGCTG 2520

Qy 560 IleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThr 579  
Db 2521 ATTACAGATGAGGAAGAAATCAAGGCTACCAACCCCGTGGCCACAGAAGAATATGGAGCA 2580  
Qy 580 ValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMet 599  
Db 2581 GTGGCCATCAACAACACGCGCCCAATACGACGCGGCGAGACCGGACTCGTGCAACACAG 2640  
Qy 600 GlyAlaLeuProGlyMetValTipGlnAspArgAspValTyrLeuGlnGlyProIleTrp 619  
Db 2641 GGGGTGATTCGCGCATGGTGTGGCAGATAGAGACGTTGTTACCTGCAGGCTCCCATCTGG 2700  
Qy 620 AlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGly 639  
Db 2701 GCCAAAATTCCTCACACGACGCGCAACTTTCACCCGTTCTCCCTGTATGGCGGCTTTGA 2760  
Qy 640 LeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProPro 659  
Db 2761 CTGAAGCACCCGCTCCTCAAATTTCTCATCAAGAACACACCGGTTCCAGCGGACCCGCG 2820  
Qy 660 AlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnVal 679  
Db 2821 CTTACCTTCAACCCAGGCAAGCTGAACCTTTTTCATCAGCAGTACAGCACCGACAGGTC 2880  
Qy 680 SerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluVal 699  
Db 2881 AGCGTGGAATCGAGTGGGAGCTGCAGAAAGAAAACAGCAAAACGCTGGAATCCAGAGATT 2940  
Qy 700 GlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGly 719  
Db 2941 CAATACACTTCCAACCTACTACAAATCTACAAATGTGGACTTTGTGTCAACACGGAAGGA 3000  
Qy 720 LeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736  
Db 3001 GTTTATAGCAGGACCTCGCCCCATTGGCACCCCGTTACCTCACCCGCAACCTG 3051

RESULT 8  
US-10-291-583-27  
; Sequence 27, Application US/10291583  
; Publication No. US20030138772A1  
; GENERAL INFORMATION:  
; APPLICANT: Gao, Guangping  
; APPLICANT: Wilson, James M.  
; APPLICANT: Alvira, Mauricio  
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus  
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1e1 Sequences Identif.  
; FILE REFERENCE: UPN-02735USA  
; CURRENT APPLICATION NUMBER: US/10/291,583  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: US 60/350,607  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/341,117  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: US 60/377,066  
; PRIOR FILING DATE: 2002-05-01  
; PRIOR APPLICATION NUMBER: US 60/386,675  
; PRIOR FILING DATE: 2002-06-05  
; NUMBER OF SEQ ID NOS: 120  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 27  
; LENGTH: 3128  
; TYPE: DNA  
; ORGANISM: 42.8  
US-10-291-583-27  
  
Alignment Scores:  
Pred. No.: 0 Length: 3128  
Score: 3453.00 Matches: 628  
Percent Similarity: 91.19% Conservative: 45  
Best Local Similarity: 85.09% Mismatches: 63  
Query Match: 86.56% Indels: 2  
DB: 13 Gaps: 2

|   |      |   |
|---|------|---|
| US-09-807-802A-13 (1-736) x US-10-291-583-27 (1-3128) |      |   |
| QY  | 1    | MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20       |
| Db  | 844  | ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTCTCTGAGGCGATTCCG 903      |
| QY  | 21   | GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp 40       |
| Db  | 904  | GAGTGGTGGGACTTGAAACCTGGAGCCCCGAAACCCAAAGCCAACCGAGCAAAAGCAGGAC 963     |
| QY  | 41   | AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp 60       |
| Db  | 964  | GACGGCCGGGTCTGGTGCTTCTGGCTACAAAGTACCTCGGACCTTCCAACGGACTCGAC 1023      |
| QY  | 61   | LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80          |
| Db  | 1024 | AAGGGGAGCCCGTCAACGGCGGCGAGCGCGGCTCGAGCACGACACAGGCTTACGAC 1083         |
| QY  | 81   | GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100      |
| Db  | 1084 | CAGCAGCTCAAAGCGGGTGACAATCCGTACCTGCGGTATAACCAACGCGCGACGCCGAGTTT 1143   |
| QY  | 101  | GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120      |
| Db  | 1144 | CAGGAGCGTCTGCAAGAGATACGTCTTTTGGGGGCAACCTCGGGGAGCAGTCTTCCAG 1203       |
| QY  | 121  | AlaLysLysArgValLeuGluProLeuGlyLeuValGluGluGlyAlaLysThrAlaPro 140      |
| Db  | 1204 | GCCAAAGAAGCGGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGCTCCT 1263     |
| QY  | 141  | GlyLysLysArgProValGluGlnSerProGln---GluProAspSerSerSerGlyIle 159      |
| Db  | 1264 | GGAAGAAGAGACCGGTAGAGCCATCACCCCGAGCGTCTCTCAGACTCCTCTACGGGCATC 1323     |
| QY  | 160  | GlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSer 179      |
| Db  | 1324 | GGCAAGACAGGCCAGCAGCGCGGAAAGAGACTCAACTTGGGCAGACTGGCGGACTCA 1383        |
| QY  | 180  | GluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGly 199      |
| Db  | 1384 | GAGTCAGTGCCCGACCCCTCAACCAATCGGAGAACCCCGCAGGCCCTCTGCTCTGGGA 1443       |
| QY  | 200  | ProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAsp 219         |
| Db  | 1444 | TCTGGTACAATGGCTGCAGCGGTGGCGCTCCAATGGCAGACAATAACGAAGGCGCCGAC 1503      |
| QY  | 220  | GlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgVal 239      |
| Db  | 1504 | GGAGTGGGTAGTTCTCTCAGAAATTTGGCATTTGCCATTCCACATGGCTGGGCGACAGATC 1563    |
| QY  | 240  | IleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGln 259      |
| Db  | 1564 | ATCACCAACAGCACCCGAACTGGGCGCTCCCACTTCCACCTACAAACCAACCACTTACAAGCAA 1623 |
| QY  | 260  | IleSerSerAlaSerThrGlyAlaSer---AsnAspAsnHisTyrPheGlyTyrSerThr 278      |
| Db  | 1624 | ATCTCCAACGGGACATCGGAGGAAGCACCAACGACACACCTACTTCGGCTACAGCACC 1683       |
| QY  | 279  | ProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGln 298      |
| Db  | 1684 | CCCTGGGGGTATTTTGACTTTAACAGATTCCACTGCCACTTCTCACCACGTGACTGGCAG 1743     |
| QY  | 299  | ArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsn 318      |
| Db  | 1744 | CGACTCATCAACAACACTGGGGATTCCGGCCCCAAGAGACTCAACTTCAAGCTCTTCAAC 1803     |
| QY  | 319  | IleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThr 338      |
| Db  | 1804 | ATCCAGGTCAAGGAGGTACCGCAGAAATGAAGGCCAACGACCATCGCCCAATAACCTTACC 1863    |
| QY  | 339  | SerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAla 358      |

---

|    |      |   |
|----|------|---|
| Db | 1864 | AGCACGATTCAGGTCTTTACGGACTCGGAATACCAGCTCCCGTACGTCCTCGGCTCTGCG 1923     |
| QY | 359  | HisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyr 378      |
| Db | 1924 | CACCAGGGCTGCCCTCCCTCCGTTCCCGCGGACGCTCTTTCATGATTCTCAGTACGGGTAC 1983    |
| QY | 379  | LeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyr 398      |
| Db | 1984 | CTGACTCTGAACAACAGCGCAGTCAGGCGGTGGGCGCTTCTCCTTCTACTGCTGGAGTAC 2043     |
| QY | 399  | PheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGlu 418      |
| Db | 2044 | TTTCTCTTCTCAAAATGCTGAGAACCGGCAACAACCTTTGAGTTTCAGCTACCACTTTGAGGAC 2103 |
| QY | 419  | ValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeu 438      |
| Db | 2104 | GTGCCCTTTTCACAGCAGCTACGCGCACAGCCAAAGCCTGGACCGGCTGATGAACCCCTC 2163     |
| QY | 439  | IleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsn 458      |
| Db | 2164 | ATCGACCAGTACCTGTACTACCTGTCTCGGACTCAGTCCACGGGAGGTACCGCAGGAACT 2223     |
| QY | 459  | LysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrp 478      |
| Db | 2224 | CAGCAGTTGCTATTCTCAGGCGCGGCTTAATAACATGTGCGTCTAGGCCCAAAAACCTGG 2283     |
| QY | 479  | LeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsn 498      |
| Db | 2284 | CTACCCGGGCGCTGTACCGGCAGCAACGCGTCTCCACGACACTGTCCGCAAAATAACAAAC 2343    |
| QY | 499  | SerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIle 518      |
| Db | 2344 | AGCAACTTTGTGGACCGGTGCCACCAAGTATCATCTGAATGGCAGAGACTCTCTGGTA 2403       |
| QY | 519  | AsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSer 538      |
| Db | 2404 | AATCCCGGTGTCGCTATGGCAACGCAAGGACGACGAAGAGCGGATTTTTCATCCAGC 2463        |
| QY | 539  | GlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnVal 558      |
| Db | 2464 | GGAGTCTTGATGTTTGGGAAACAGGGGAGCTGGAAAAGACAACGCTGGACTATAGCAGCGTT 2523   |
| QY | 559  | MetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGly 578      |
| Db | 2524 | ATGCTAACCAAGTAGGAAGAAATCAAAACCAACCAACCCAGTGGCCACAGAACAGTACGGC 2583    |
| QY | 579  | ThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAla 598      |
| Db | 2584 | GTGGTGGCCGATAAACCCTGCNAACAGCAAAACGCCGCTCCTATTGTAGGGGCGTCAACAGT 2643   |
| QY | 599  | MetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIle 618      |
| Db | 2644 | CAAGGAGCCTTACCTGGCATGGTCTGSCAGAACCGGACGCTGTACCTGCAGGGTCTTATC 2703     |
| QY | 619  | TrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPhe 638      |
| Db | 2704 | TGGGCCAAGATTCCTCACCGGACGGCAACTTTCATCTCTCGCCGCTGATGGAGGCTTT 2763       |
| QY | 639  | GlyLeuLysAsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsnPro 658      |
| Db | 2764 | GGACTGAAACACCCGCCCTCTCAGATCTGATTAAGAATACACCTGTTCGCGGATCCT 2823        |
| QY | 659  | ProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGln 678      |
| Db | 2824 | CCAACTACCTTCAGTCAAGCCCAAGTGGCGTCTGTTTCACTACGTCAGTACAGCACCGGACAG 2883  |
| QY | 679  | ValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGlu 698      |
| Db | 2884 | GTCAGCGTGGAATTTGAATGGAGCTGCAGAAAGAGAACAGCAAGCGCTGGAACCCAGAG 2943      |
| QY | 699  | ValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsn 718      |
| Db | 2944 | ATTCAGTATACTCCAACTACTACAAATCTACAAATGTGGACTTTTGTCTGTCATAATACTGAG 3003  |



```
Qy 719 GlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736
    |||:::|||||
Db 3004 GGTACTATTACAGAGCCTCGCCCATGGCACCCGTTACCTCACCCGTAACCTG 3057

RESULT 9
US-10-291-583-26
; Sequence 26, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 3075
; TYPE: DNA
; ORGANISM: new AAV serotype, clone H2
US-10-291-583-26

Alignment Scores:
Pred. No.: 0 Length: 3075
Score: 3451.50 Matches: 622
Percent Similarity: 91.17% Conservative: 49
Best Local Similarity: 84.51% Mismatches: 64
Query Match: 86.53% Indels: 1
DB: 13 Gaps: 1

US-09-807-802A-13 (1-736) x US-10-291-583-26 (1-3075)

Qy 1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20
    |||:::|||||
Db 803 ATGGCTGCCGATGGTTATCTCCAGATTGGCTCGAGGACACTCTCTCTGAAGGGATAAGA 862

Qy 21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp 40
    :::::|||||
Db 863 CAGTGGTGAAGCTCAAACTGGCCCAACCACCACCAAGCCCGCAGAGCGGCATAAGGAC 922

Qy 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp 60
    |||:::|||||
Db 923 GACAGCAGGGGTCTGTGTCTTCTGGGTACAAGTACCTCGGACCCCTCAACGGACTCGAC 982

Qy 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80
    |||:::|||||
Db 983 AAGGGGAGCCGTCAACGAGGACAGCCGCGGCCCTCGAGCAGCACAAGGCCTACGAC 1042

Qy 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100
    :::|||||
Db 1043 CGGCAGCTCGACAGCGGAGACAACCCGTTACTCAAGTACAACCAACCGCGCAGAGTTT 1102

Qy 101 GlnGluArgLeuGlnGluAspThrSerPheGlyAsnLeuGlyArgAlaValPheGln 120
    |||:::|||||
Db 1103 CAGGAGCGCCTTAAAGAAGATACGTCTTTTGGGGGCAACCTCGGACGAGCAGTCTTCCAG 1162

Qy 121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro 140
    |||:::|||||
Db 1163 GCGAAAAGAGGGTTCTTGAACCTCTGGGCCTGGTTGAGGAACCTGTAAAGACGGCTCCG 1222

Qy 141 GlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerGlyIleGly 160
    |||:::|||||
```

```
Db 1223 GGAAAAAGAGCGGTAGAGCACTCTCTGTGGAGCCAGACTCTCTCTCGGAACCGGA 1282
    |||:::|||||
Qy 161 LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180
    |||:::|||||
Db 1283 AAAGCGGGCCAGCGGCTGCAAGAAAAAGATTAAATTTTGGTCAGACTGGAGACGAC 1342
    |||:::|||||
Qy 181 SerValProAspProGlnProLeuGlyGluProAlaThrProAlaAlaValGlyPro 200
    |||:::|||||
Db 1343 TCCGTACCTGACCCCGCAGCTCTCGAGACGCCACCGAGCCCTCTCTGGTCTGGGATCT 1402
    |||:::|||||
Qy 201 ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnGluGlyAlaAspGly 220
    |||:::|||||
Db 1403 ACTACAATGGCTACAGCGAGTGGCGCACCAATGGCAGACAATAACGAGGGTGCGGATGGA 1462
    |||:::|||||
Qy 221 ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIle 240
    |||:::|||||
Db 1463 GTGGGTAAATTCCTCAGGAATTTGGCATTTGCCATTTCCCAATGGCTGGCGACAGATCATC 1522
    |||:::|||||
Qy 241 ThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIle 260
    |||:::|||||
Db 1523 ACCACGACGCCCGAACCTGGGCCCTGCCACATACAAACCAACCTCTACAAAGCAATC 1582
    |||:::|||||
Qy 261 SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp 280
    |||:::|||||
Db 1583 TCCAGCCAATCA---GGAGCCAGCAACGACACCACTACTTTTGGCTACAGCACCCCTGG 1639
    |||:::|||||
Qy 281 GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu 300
    |||:::|||||
Db 1640 GGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCCACCACGCTGCTGCAAGACTC 1699
    |||:::|||||
Qy 301 IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln 320
    |||:::|||||
Db 1700 ATCAACAACAACCTGGGATTTCCGGCCCCAAAGACTCAACTTCAAGCTCTTTAATATCAA 1759
    |||:::|||||
Qy 321 ValLysGluValThrThrAsnAspGlyValThrIleAlaAsnAsnLeuThrSerThr 340
    |||:::|||||
Db 1760 GTCAAGAGAGGTACGCAAGATGACGGTACGACGACGATGGCCAATAACCTTACCAGCAG 1819
    |||:::|||||
Qy 341 ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln 360
    |||:::|||||
Db 1820 GTTCAGGTGTTTACTGACTCGGAGTACCAGTCCCCGTACGCTCTGGGCTCGGCGCATCAA 1879
    |||:::|||||
Qy 361 GlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr 380
    |||:::|||||
Db 1880 GGATGCTCTCCGCGCGTTTCCAGCGGACGCTCTTCTATGTTGCCACAGTATGGATACCTCAC 1939
    |||:::|||||
Qy 381 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 400
    |||:::|||||
Db 1940 CTGAACAACGGGAGTACGCGGTAGGACGCTCTTCTCTTTTACTGCTGGAGTACTTCTCT 1999
    |||:::|||||
Qy 401 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro 420
    |||:::|||||
Db 2000 TCTCAGATGCTGCGTACTGGAAACAACCTTTCAGTTTCAGTTCAGTACACATTTTGAAGAC 2059
    |||:::|||||
Qy 421 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 440
    |||:::|||||
Db 2060 TTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGATCGGCTGATGAATCCTCTGTATCGAC 2119
    |||:::|||||
Qy 441 GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAsp 460
    |||:::|||||
Db 2120 CAGTACCTGTATTATCTGAACAAGACACAAATAAGTGGAACTCTTCAGCAGTCTCGG 2179
    |||:::|||||
Qy 461 LeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuPro 480
    |||:::|||||
Db 2180 CTACTGTTTAGCCAAGCTGGACCAACCAACATGTCTCTCAAGCTAAAAAAGTGGCTGCT 2239
    |||:::|||||
Qy 481 GlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnSerAsn 500
    |||:::|||||
Db 2240 GGACTTGTCTACAGACAGCAGCGGTCTGTCAAAACAGGCAAAACGACAAACACACAGCAAC 2299
    |||:::|||||
Qy 501 PheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnPro 520
    |||:::|||||
```

Db 2300 TTTCCCTGGACTGCAGCTACAAAGTATCATCTAAATGGCCGGGACTCGTTGGTTAATCCA 2359

Qy 521 GlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyVal 540  
|||||

Db 2360 GGACCAGCTATGGCCAGTCACAAGGATGACGAAGAAAGTTTTCCTCCCATGCATGGAACC 2419

Qy 541 MetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIle 560  
:::|||||

Db 2420 CTGATATTGGTAAACAAGGAACAAATGCCAACGACGGGATTGTGAAAATGTCTATGATT 2479

Qy 561 ThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrVal 580  
|||||

Db 2480 ACAGATGAAGAAGAAATCAGGGCCACCAATCCCGTGGCTACGGAGCAGTACGGGACTGTG 2539

Qy 581 AlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGly 600  
:::|||||

Db 2540 TCAAATAATTGCAAACTCAACACTGGTCCCACTACTGGAAGTGTCAATCGCCAAGGA 2599

Qy 601 AlaLeuProGlyMetValTyrGlnAspArgAspValTyrLeuGlnGlyProIleTyrAla 620  
|||||

Db 2600 GCGTTACCTGGTATGGTGGCAGGATCGAGACGTGTACCTGCAGGGACCCCAATTGGGCC 2659

Qy 621 LysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeu 640  
|||||

Db 2660 AAGATTCTCACACCGATGGACACTTTCATCTTCTCCACTGATGGAGGTTTTGGACTC 2719

Qy 641 LysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAla 660  
:::|||||

Db 2720 AAACACCGCCTCCTCAGATCATGATCAAAAACACTCCCGTTCAGGCCAATCCTCCAC 2779

Qy 661 GluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSer 680  
|||||

Db 2780 AACTTTCAGTTCTGCCAAGTTTGTCTTTCATCACACAGTATTCACGGGACAGTTCAGC 2839

Qy 681 ValGluIleGluTyrGluLeuGlnLysGluAsnSerLysArgTyrAsnProGluValGln 700  
|||||

Db 2840 GTGGAGATCGAGTGGGAGCTGCAGAAGGAGAACACGCAACGCTGGAATCCCGAATTCAG 2899

Qy 701 TyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeu 720  
|||||

Db 2900 TACACTTCCAACACTACAACAAGTCTGTTAATGTGGACTTTACTGTGGACACTAATGGTGTG 2959

Qy 721 TyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736  
|||||

Db 2960 TATTCAGAGCCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTG 3007

RESULT 10

US-10-291-583-59

; Sequence 59, Application US/10291583

; Publication No. US20030138772A1

; GENERAL INFORMATION:

; APPLICANT: Gao, Guangping

; APPLICANT: Wilson, James M.

; APPLICANT: Alvira, Mauricio

; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus (

; FILE REFERENCE: UPN-02735USA

; CURRENT APPLICATION NUMBER: US/10/291,583

; CURRENT FILING DATE: 2002-11-12

; PRIOR APPLICATION NUMBER: US 60/350,607

; PRIOR FILING DATE: 2001-11-13

; PRIOR APPLICATION NUMBER: US 60/341,117

; PRIOR FILING DATE: 2001-12-17

; PRIOR APPLICATION NUMBER: US 60/377,066

; PRIOR FILING DATE: 2002-05-01

; PRIOR APPLICATION NUMBER: US 60/386,675

; PRIOR FILING DATE: 2002-06-05

; NUMBER OF SEQ ID NOS: 120

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 59

; LENGTH: 3129

; TYPE: DNA

; ORGANISM: 44.2

US-10-291-583-59

Alignment Scores:

Pred. No.: 0 Length: 3129

Score: 3448.00 Matches: 627

Percent Similarity: 91.19% Conservative: 46

Best Local Similarity: 84.96% Mismatches: 63

Query Match: 86.44% Indels: 2

DB: 13 Gaps: 2

US-09-807-802A-13 (1-736) x US-10-291-583-59 (1-3129)

Qy 1 MetAlaAlaAspGlyTyrLeuProAspTyrLeuGluAspAsnLeuSerGluGlyIleArg 20  
|||||

Db 845 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTCGC 904

Qy 21 GluTyrTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnLysGlnAsp 40  
|||||

Db 905 GAGTGGTGGGACTTGAAACCTGGAGCCCCGAAACCCAAAGCCACCAAGCAAGCAGGAC 964

Qy 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp 60  
|||||

Db 965 GACGGCCGGGTCTGGTCTTCTGGCTACAAGTACCTCGGACCTTCAACGGACTCGAC 1024

Qy 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80  
|||||

Db 1025 AAGGGGAGCCCGTCAACGCGGGACGAGCGGGCCCTCGAGCACGACAAGGCCTACGAC 1084

Qy 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100  
|||||

Db 1085 CAGCAGCTCAAAAGCGGGTGACAAATCCGTACCTGCGGTATTAACCAACGCGCAGCGCGAGTTT 1144

Qy 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120  
|||||

Db 1145 CAGGAGCGTCTGCAAGAAGATACGTCTTTTGGGGCAACCTCGGGCGAGCAGTCTTCCAG 1204

Qy 121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro 140  
|||||

Db 1205 GCCAAGAAGCGGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGGCGCTAAGACGGCTCCT 1264

Qy 141 GlyLysLysArgProValGluGlnSerProGln--GluProAspSerSerSerGlyIle 159  
|||||

Db 1265 GGAAGAAGAGACCGGTAGAGCCATCACCCAGCGGTTCTCCAGACTCTCTACGGGCATC 1324

Qy 160 GlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSer 179  
|||||

Db 1325 GGCNAGAAAGGCCAGCAGCCCCGCGAAAAAGAGACTCAACTTTGGGCAGACTGGCGACTCA 1384

Qy 180 GluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGly 199  
|||||

Db 1385 GAGTCAGTCCCGACCCCTCAACCAATCGAGAAACCCCGCAGGCCCTCTGGTCTGGGA 1444

Qy 200 ProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAsp 219  
|||||

Db 1445 TCTGGTACAATGGCTGCAGGCGGTGGCGTCCAATGGCAGACAATAACGAAGGCCCGAC 1504

Qy 220 GlyValGlyAsnAlaSerGlyAsnTyrHisCysAspSerThrTyrLeuGlyAspArgVal 239  
|||||

Db 1505 GGAGTGGGTAGTCTCTCAGGAAATTGGCATTTCCGATTCCACATGGCTGGCGCAGAGTC 1564

Qy 240 IleThrThrSerThrArgThrTyrAlaLeuProThrTyrAsnAsnHisLeuTyrLysGln 259  
|||||

Db 1565 ATCACCACGACCCGAAACCTGGGCCCTCCACACCTACAACACCACTCTACAAAGCAA 1624

Qy 260 IleSerSerAlaSerThrGlyAlaSer---AsnAspAsnHisTyrPheGlyTyrSerThr 278  
|||||

Db 1625 ATCTCCAACGGGACTTCGGGAGGAAGCACCACCAACGACAAACACCTACTTCCGCTACAGCAC 1684

Qy 279 ProTyrGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTyrGln 298  
|||||

Db 1685 CCCTGGGGGTATTTGACTTTAACAGATTCCACTGCCACTTCTCACCACCGTACTGGCAG 1744

Qy 299 ArgLeuIleAsnAsnAsnTyrGlyPheArgProLysArgLeuAsnPheLysLeuPheAsn 318

Db 1745 CGACTCATCAACAACAACACTGGGGAATTCGGGCCCAAGAGACTCAACTTCAAGCTCTTCAAC 1804

Qy 319 ILeGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThr 338

Db 1805 ATCCAGGTCAGAGAGGTACGCGAGAATGAAGGCACCAAGACCATCGCCAATAACCTTACC 1864

Qy 339 SerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAla 358

Db 1865 AGCAGGATTCAGGTCTTTACGGACTCGGAATACCAGCTCCCGTACGTCTCTCGGCTCTGCG 1924

Qy 359 HisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyr 378

Db 1925 CACCAGGGCTGCCTGCCTCCGTTCCCGGGGACGCTCTTCATGATTCCTCAGTACGGGTAC 1984

Qy 379 LeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyr 398

Db 1985 CTGACTCTGAACAATGGCAGTCAGGCCGTGGGCCGTTCTCTCTACTGCCTGGAGTAC 2044

Qy 399 PheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGlu 418

Db 2045 TTTCTCTCTCAAAATGCTGAGACGGGCAACAACCTTTGAGTTTCAGTACCAGTTTGAGGAC 2104

Qy 419 ValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeu 438

Db 2105 GTGCCCTTTTCAGCAGCTACGCGCACAGCCAAAGCCTGGACCGGTGATGAACCCCTC 2164

Qy 439 IleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsn 458

Db 2165 ATCAGCACGATCTGTACTACCTGTCTCGGACTCAGTCCACGGGAGGTACCGCAGGAAC 2224

Qy 459 LysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTyr 478

Db 2225 CAGCAGTTTGCTATTTTCTCAGGCCGGGCGCTTAATAACATGTGGCTCAGGCCAAAACTGG 2284

Qy 479 LeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 498

Db 2285 CTACCCGGGCCCTGCTACCGGCAGCAACGGCTCTCCACGACACTGTCCGCAAAATAACAAC 2344

Qy 499 SerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIle 518

Db 2345 AGCAACTTTGCCTGGACCGGTGCCACCAGTATCATCTGAATGGCAGAGACTCTCTGGTA 2404

Qy 519 AsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSer 538

Db 2405 AATCCCGGTGTCGCTATGGCAACCCACAAAGGACGACGAAGCGATTTTTCGTCCAGC 2464

Qy 539 GlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnVal 558

Db 2465 GGAGTCTTAATGTTTGGGAAACAGGGAGCTGGAAAAGACAACGTGGACTATAGCAGCGTT 2524

Qy 559 MetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGly 578

Db 2525 ATGCTAACCACTGAGGAAGAAATTAACCAACCAACCCAGTGGCCACAGAACAGTACGGC 2584

Qy 579 ThrValAlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAla 598

Db 2585 GTGGTGGCCGATAACCTGCAACAGCAAAACGCCGCTCCTATTGTAGGGGCCGTCAACAGT 2644

Qy 599 MetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIle 618

Db 2645 CAAGGAGCCTTACCTGGCATGGTCTGGCAGAACCGGACGTGTACCTGCAAGGTCTTATC 2704

Qy 619 TrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPhe 638

Db 2705 TGGGCCAAGATTCTCTACACGGACGGAAACTTTTCATCCCTCGCCGCTGATGGAGGCTTT 2764

Qy 639 GlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnPro 658

Db 2765 GGACTGAAACACCCCGCTCTCAGATCCTGTATTAGAAATACACCTGTTCCCGCGATCCT 2824

Qy 659 ProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGln 678

Db 2825 CCAACTACCTTCAGTCAAGCTAAGCTGGCGTTCGTTCATCAGCAGTACAGCACCGGACAG 2884

Qy 679 ValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGlu 698

Db 2885 GTCAGCGTGGAAATGAATGGGAGCTGCAGAAAGAAACAGCAACGCTGGAAACCCAGAG 2944

Qy 699 ValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsn 718

Db 2945 ATTCAATACACTTCCAACACTACTACAAATCTACAAATGTGGACTTTTGTCTGTTAACACAGAT 3004

Qy 719 GlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736

Db 3005 GGCACCTATTCTGAGCCTCGCCCCATCGGCACCCGTTACCTCACCCTAATCTG 3058

RESULT 11

US-10-291-583-25/c

; Sequence 25, Application US/10291583

; Publication No. US20030138772A1

; GENERAL INFORMATION:

; APPLICANT: Gao, Guangping

; APPLICANT: Wilson, James M.

; APPLICANT: Alvira, Mauricio

; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus

; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1e1 Sequences Identif

; FILE REFERENCE: UPN-02735USA

; CURRENT APPLICATION NUMBER: US/10/291,583

; CURRENT FILING DATE: 2002-11-12

; PRIOR APPLICATION NUMBER: US 60/350,607

; PRIOR FILING DATE: 2001-11-13

; PRIOR APPLICATION NUMBER: US 60/341,117

; PRIOR FILING DATE: 2001-12-17

; PRIOR APPLICATION NUMBER: US 60/377,066

; PRIOR FILING DATE: 2002-05-01

; PRIOR APPLICATION NUMBER: US 60/386,675

; PRIOR FILING DATE: 2002-06-05

; NUMBER OF SEQ ID NOS: 120

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 25

; LENGTH: 3142

; TYPE: DNA

; ORGANISM: new AAV serotype, clone H6

US-10-291-583-25

Alignment Scores:

Pred. No.: 0 Length: 3142

Score: 3446.50 Matches: 622

Percent Similarity: 91.17% Conservative: 49

Best Local Similarity: 84.51% Mismatches: 64

Query Match: 86.40% Indels: 1

DB: 13 Gaps: 1

US-09-807-802A-13 (1-736) x US-10-291-583-25 (1-3142)

Qy 1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20

Db 2339 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACACTCTCTCTGAAGGAATAAGA 2280

Qy 21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnLysGlnAsp 40

Db 2279 CAGTGGTGGAAAGCTCAAAACCTGGCCCAACCAACCAAGCCCGCAGAGCGGCATTAAGGAC 2220

Qy 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyPropheAsnGlyLeuAsp 60

Db 2219 GACAGCAGGGGTCTTGTGCTTCCTGGGTACAAGTACCTCGGACCTTCAACGGGACTCGAC 2160

Qy 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80

Db 2159 AAGGGGAGCCGCTCAACGAGGACGACGCCCGGCCCTCGAGCAGCAACGACCAAGGCCTACGAC 2100

Qy 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100

Db 2099 CGGCAGCTCGACAGCGGAGACAAACCCGTACCTCAAGTACAAACCAACCGCCGACGAGATTT 2040



Qy 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120  
Dbb 2039 CAGAGCGCCTTAAAGAGATACGCTCTTTGGGGCACTCGGACGAGCAGTCTTCAG 1980  
Qy 121 AlaValLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro 140  
Dbb 1979 GCGAAAAAGAGGGTTCTTGAACCTCTGGGCCTGGTGTGAGAGCCTGTTAAGACGGCTCCG 1920  
Qy 141 GlyValLysArgProValGluGlnSerProGlnGluProAspSerSerGlyIleGly 160  
Dbb 1919 GGAATAAAGAGCGCGGTAGAGCACTCTCCTGTGGAGCCAGACTCTCCTCGGGAACCGGA 1860  
Qy 161 LysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGlu 180  
Dbb 1859 AAAGCGGCCAGCAGCCCTGCAAGAAAAGATTAAATTTGGTCAGACTGGAGACCGCAGAC 1800  
Qy 181 SerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGlyPro 200  
Dbb 1799 TCCGTACCTGACCCCGAGCCTCTCGGACAGCCACCGAGCCCCCTCTGGTCTGGGATCT 1740  
Qy 201 ThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGly 220  
Dbb 1739 ACTCAATGGCTACAGCGAGTGGCGCACCAATGGCAGACAATAACGAGGGTGGCGATGGA 1680  
Qy 221 ValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValile 240  
Dbb 1679 GTGGGTAATTCCTCAGGAAATTTGGCATTGGCATTTCCTCAATGGCTGGCGACAGAGTCATC 1620  
Qy 241 ThrThrSerThrArgThrTrpAlaLeuProThrTrpAsnAsnHisLeuTyrLysGlnIle 260  
Dbb 1619 ACCACAGCACCCGAACCTGGGCCCTGCCACATACAAACAACCACTCTACAAGCAAATC 1560  
Qy 261 SerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrp 280  
Dbb 1559 TCCAGCCAATCA--GGAGCCAGCAACGACAACCACTACTTTGGCTACAGCACCCCTGG 1503  
Qy 281 GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu 300  
Dbb 1502 GGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCACCACGTGACTGGCAAGACTC 1443  
Qy 301 IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln 320  
Dbb 1442 GTCAACAACAACCGGGATTCCGGCCCCAAAAGACTCAACTTCAAGCTCTTTAATATTCAA 1383  
Qy 321 ValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThr 340  
Dbb 1382 GTCAAGAGGTCACGCAGATGACGGTACGACGACGATTGCCAATAACCTTACCAGCACG 1323  
Qy 341 ValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGln 360  
Dbb 1322 GTTCAGTGTTTACTGACTCGGAGTACCAGCTCCCGTAGCTCCTGGCTCGGGCGCATCAA 1263  
Qy 361 GlyCysLeuProPropheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThr 380  
Dbb 1262 GGATGCCTCCCGCGTTTCCAGCGGACGCTTTCATGGTGGTCCACAGTATGGATACCTCACC 1203  
Qy 381 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 400  
Dbb 1202 CTGAACAACGGGAGTCAGCGGTAGGACGCTCTTCTCTTTACTGCCTGGAGTACTTTCCT 1143  
Qy 401 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro 420  
Dbb 1142 TCTCAGATGCTGCGTACTGGAAACAACCTTTTCAGTTCAGTACACTTTTGAAGACGCTGCCT 1083  
Qy 421 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 440  
Dbb 1082 TTCCACAGCAGCTACGCTACAGCCAGAGTCTGGATCGGCTGATGAATCCTCTGATCGAC 1023  
Qy 441 GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAsp 460  
Dbb 1022 CAGTACCTGTATTATCTGAACACAGACACAACAATAATAGTGAAGTCTTTCAGCAGTCTCGG 963  
Qy 461 LeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuPro 480

Dbb 962 CTACTGTTTAGCCCAAGCTGGACCAACCAACATGTCTCTTCAAGCTAAAAAAGTGGTGCCT 903  
Qy 481 GlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnSerAsn 500  
Dbb 902 GGACCTTGTACAGACAGCAGCGTCTGTCAAAACAGGCAACAGACAACAACAGCAAC 843  
Qy 501 PheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnPro 520  
Dbb 842 TTTCCCTGGACTGCAGCTACAAAGTATCATCTAAATGGCCGGGACTCGTTGGTTAATCCA 783  
Qy 521 GlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyVal 540  
Dbb 782 GGACCAGCTATGGCCAGTCACAAGGATGACGAAGAAAAGTTTTTCCCATGCATGGAACC 723  
Qy 541 MetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAsnValMetile 560  
Dbb 722 CTGATATTTGGTAAACAAGGAACAATAATGCCAACGACGCGGATTTGGAAAATGTCATGATT 663  
Qy 561 ThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrVal 580  
Dbb 662 ACAGATGAAGAAGAAATCAGGGCCACCAATCCCGTGGTACGAGCAGTACGGGACTGTG 603  
Qy 581 AlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGly 600  
Dbb 602 TCAATAATTTGCAAAACACTCAAAACACTGGTCCAACTACTGGAACGTGTCAATCACCAAGGA 543  
Qy 601 AlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAla 620  
Dbb 542 GCGTTACCTGGTATGGTGTGGCAGGATCGAGACGTGTACCTGCAGGGAGCCATTTGGGCC 483  
Qy 621 LysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeu 640  
Dbb 482 AAGATTCTCACACCGATGGACACTTTTCATCTCTTCCACTGATGGGAGGTTTGGACTC 423  
Qy 641 LysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAla 660  
Dbb 422 AAACACCCCGCTCCTCAGATCATGATCAAAAACACTCCCGTTCCAGCCAATCCTCCACA 363  
Qy 661 GluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSer 680  
Dbb 362 AACTTCAGTCTCGCAAGTTTGCTTCTTTCATCACACAGTATTCACCGGACAGGTGAGC 303  
Qy 681 ValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGln 700  
Dbb 302 GTGGAGATCGAGTGGGAGCTGCAGAGAGGAGAACAGCAACGCTGGAAATCCCGAAATTCAG 243  
Qy 701 TyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeu 720  
Dbb 242 TACACTTCCAACATAACAAGTCTGTTAATGTGGACTTTTACTGTGGACACTAATGTTGTG 183  
Qy 721 TyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736  
Dbb 182 TATTTCAGAGCCTCGCCCCATTGGCACCATGATACCTGACTCGTAATCTG 135

RESULT 12

US-10-291-583-41

; Sequence 41, Application US/10291583

; Publication No. US20030138772A1

; GENERAL INFORMATION:

; APPLICANT: Gao, Guangping

; APPLICANT: Wilson, James M.

; APPLICANT: Alvira, Mauricio

; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus

; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1el Sequences Identif:

; FILE REFERENCE: UPN-02735USA

; CURRENT APPLICATION NUMBER: US/10/291,583

; CURRENT FILING DATE: 2002-11-12

; PRIOR APPLICATION NUMBER: US 60/350,607

; PRIOR FILING DATE: 2001-11-13

; PRIOR APPLICATION NUMBER: US 60/341,117

; PRIOR FILING DATE: 2001-12-17

; PRIOR APPLICATION NUMBER: US 60/377,066

; PRIOR FILING DATE: 2002-05-01  
; PRIOR APPLICATION NUMBER: US 60/386,675  
; PRIOR FILING DATE: 2002-06-05  
; NUMBER OF SEQ ID NOS: 120  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 41  
; LENGTH: 3123  
; TYPE: DNA  
; ORGANISM: 43.12  
US-10-291-583-41

Alignment Scores:

|                        |         |               |      |
|------------------------|---------|---------------|------|
| Pred. No.:             | 0       | Length:       | 3123 |
| Score:                 | 3445.00 | Matches:      | 625  |
| Percent Similarity:    | 91.19%  | Conservative: | 48   |
| Best Local Similarity: | 84.69%  | Mismatches:   | 63   |
| Query Match:           | 86.36%  | Indels:       | 2    |
| DB:                    | 13      | Gaps:         | 2    |

US-09-807-802A-13 (1-736) x US-10-291-583-41 (1-3123)

|    |      |   |      |
|----|------|---|------|
| QY | 1    | MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg    | 20   |
| Db | 839  | ATGGCTGCCGATGGTTATCTCCAGATTGGCTTGAGGACAACCTCTCTGAGGGCAATTCGC    | 898  |
| QY | 21   | GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnLysGlnAsp       | 40   |
| Db | 899  | GAGTGGTGGACCTGAACCTGGAGCCCCGGAACCCAAAGCCCAACAGCAAGCAGGAC        | 958  |
| QY | 41   | AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp    | 60   |
| Db | 959  | GACGGCCGGGTCTGGTGTCTCTGGCTACAAGTACCTCGGACCTTCAACGGACTCGAC       | 1018 |
| QY | 61   | LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp       | 80   |
| Db | 1019 | AAGGGGAGCCCGTCAACGCGCGGACGCGCGCCCTCGAGCACGACAAGGCCCTACGAC       | 1078 |
| QY | 81   | GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe    | 100  |
| Db | 1079 | CAGCAGCTCAAGCGGGTGACAAATCCGTACCTCGGTATATAACCAACGCGACGCCGAGTTT   | 1138 |
| QY | 101  | GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln    | 120  |
| Db | 1139 | CAGGAGCGTCTGCAAGAAGATACGTCTTTTGGGGCAACCTCGGGCGAGCAGTCTTCCAG     | 1198 |
| QY | 121  | AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro       | 140  |
| Db | 1199 | GCCAAAGAAGCGGGTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGCTCCT     | 1258 |
| QY | 141  | GlyLysLysArgProValGluGlnSerProGln---GluProAspSerSerSerGlyIle    | 159  |
| Db | 1259 | GGAAAGAAGAGACCGGTAGAGCCATCACCTCAGCGTTCCCCCGACTCTCTCCACGGGCATC   | 1318 |
| QY | 160  | GlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSer    | 179  |
| Db | 1319 | GGCAAGAAAGGCCACCGCCGAGCCCGGAGAAAGAGACTGAACTTTGGGCAGACTGGCGACTCG | 1378 |
| QY | 180  | GluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGly    | 199  |
| Db | 1379 | GAGTCAGTCCCCGACCTCAACCAATCGGAGAACCCACAGAGCCCTCTGGTCTGGGA        | 1438 |
| QY | 200  | ProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAsp       | 219  |
| Db | 1439 | TCTGGTACAATGGCTGCAGGCGGTGGCGTCCAAATGGCAGACAAATAACGAAGGCCCGAC    | 1498 |
| QY | 220  | GlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgVal    | 239  |
| Db | 1499 | GGAGTGGGTAGTTCCTCAGGAAATTGGCATTGGCATTCACATGGCTGGCGCAGAGTTC      | 1558 |
| QY | 240  | IleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGln    | 259  |
| Db | 1559 | ATCACCACCGACCCGAAACCTGGGCCCTGCCACCTACAACAACCATCTCTACAAGCAA      | 1618 |

|    |      |  |      |
|----|------|--|------|
| QY | 260  | IleSerSerAlaSerThrGlyAlaSer---AsnAspAsnHisTyrPheGlyTyrSerThr       | 278  |
| Db | 1619 | ATCTCCAAACGGGACATCGGAGGAAGCACTAACGACAACACCTACTTTGGCTACAGCACC       | 1678 |
| QY | 279  | ProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGln       | 298  |
| Db | 1679 | CCCTGGGGGTATTTGACTTCAACAGATTCCACTGCCACTTCTCACCACGCTGACTGGCAG       | 1738 |
| QY | 299  | ArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsn       | 318  |
| Db | 1739 | CGACTCATCAACAATAACTGGGGATTCCGGGCCCAAGAGACTCAACTTCAAGCTCTTCAAC      | 1798 |
| QY | 319  | IleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThr       | 338  |
| Db | 1799 | ATCCAGGTCAAGGAGGTACAGCAGAAATGAAGGCACCAAGACCATCGCCAAATAACCTTACC     | 1858 |
| QY | 339  | SerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAla       | 358  |
| Db | 1859 | AGCAGATTACGGTGTACGGACTCGGAATACCGACTCCCGTACGCTCCGCTCGGCTCTGCG       | 1918 |
| QY | 359  | HisGlnGlyCysLeuProPheProPheProAlaAspValPheMetIleProGlnTyrGlyTyr    | 378  |
| Db | 1919 | CACCAGGGCTGCCCTCCCTCCGTTCCCGGGCGGACGTCTTCATGATTCTCTCAGTACGGGTAT    | 1978 |
| QY | 379  | LeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyr       | 398  |
| Db | 1979 | CTGACCTTAAACAATGGCAGTCAAGGCTGTGGCGCTTCTCTTCTACTGCTTGGGAATAC        | 2038 |
| QY | 399  | PheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGlu       | 418  |
| Db | 2039 | TTCCCTTCTCAAATGCTGAGGACGGGCAACAACATTGAATTGAAATTCAGCTACACCTTCGAGGAC | 2098 |
| QY | 419  | ValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeu       | 438  |
| Db | 2099 | GTGCTTTCCACAGCAGCTACGCGCACAGCCAGAGCCTGGACCGGTGATGATGAACCTCTC       | 2158 |
| QY | 439  | IleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsn       | 458  |
| Db | 2159 | ATCGACCACTACCTGTATTACTTATCCAGAACTCAGTCCACAGGAGGAACCTCAAGGTACT      | 2218 |
| QY | 459  | LysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrp       | 478  |
| Db | 2219 | CAGCAATTGTTATTTCTCAAGCCGGCCGCCCAACATGTGGCTCAGGCCCAAGAACTGG         | 2278 |
| QY | 479  | LeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsn       | 498  |
| Db | 2279 | CTACCTGGACCGTGTACCGTCAGCAACGAGTTTCCACGACACTGTTCGCAAAACACAAC        | 2338 |
| QY | 499  | SerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIle       | 518  |
| Db | 2339 | AGCAATTTGCTTGGACCGGTGCCACCAAGTATCACCTGAATGGCAGAGACTCCTGGTT         | 2398 |
| QY | 519  | AsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSer       | 538  |
| Db | 2399 | AATCCCGCGGTTGCCATGGCTACCCCAAGGACGACGAGGAGCGCTTCTTCCCGTCAAGC        | 2458 |
| QY | 539  | GlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnVal       | 558  |
| Db | 2459 | GGAGTCTTAATGTTGGCAAGCAGGGGGCTGGAAGAAACAATGTGGACTACAGCAGCGTG        | 2518 |
| QY | 559  | MetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGly          | 578  |
| Db | 2519 | ATGCTCACCGCAAGAAATAATAAACTACTAACCCAGTGGCTACAGAGCAGTATGGT           | 2578 |
| QY | 579  | ThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAla       | 598  |
| Db | 2579 | GTGGTGGCAGACAACCTGCAGCAGACCAACCGGAGCTCCCATTTGGGAACCTGTCAACAGC      | 2638 |
| QY | 599  | MetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIle       | 618  |
| Db | 2639 | CAGGGGCCCTTACCTGGTATGGTCTGGCAAAACCCGGGACGTGTACCTGCAGGGCCCCATC      | 2698 |
| QY | 619  | TrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPhe       | 638  |

```
Db 2699 TGGGCCAAATTCCTCACGGACGGCAACTTTCATCCTTCGCCGCTGATGGGAGCTTT 2758
Qy 639 GlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnPro 658
Db 2759 GGACTGAAACACCCGCCCTCCTCAGATCCTGGTGAAAAACACTCCTGTTCCTGCGGATCCT 2818
Qy 659 ProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGln 678
Db 2819 CCGACCACCTTCAGCCAGGCCAAGCTGGCTTCTTTTATCAGCAGTACAGCACCGGACAG 2878
Qy 679 ValSerValGluIleGluTyrGluLeuGlnLysGluAsnSerLysArgTyrAsnProGlu 698
Db 2879 GTCAGCGTGGAAATCGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAACCCAGAG 2938
Qy 699 ValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsn 718
Db 2939 ATTCAGTATACTTCCAACACTACTACAAATCTACAAATGTGGACTTTGCTGTCAATACTGAG 2998
Qy 719 GlyLeuTyrThrGluProArgProIleGlyThrArgTyrIleuThrArgProLeu 736
Db 2999 GGTACTATTACAGAGCCTCGCCCATGGCACTCGTTATCTCACCCGTAATCTG 3052

RESULT 13
US-10-291-583-1
; Sequence 1, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus
; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1el Sequences Identifi
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/350,607
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4721
; TYPE: DNA
; ORGANISM: adeno-associated virus serotype 7
US-10-291-583-1

Alignment Scores:
Pred. No.: 0 Length: 4721
Score: 3442.50 Matches: 630
Percent Similarity: 92.02% Conservative: 50
Best Local Similarity: 85.25% Mismatches: 54
Query Match: 86.30% Indels: 5
DB: 13 Gaps: 4

US-09-807-802A-13 (1-736) x US-10-291-583-1 (1-4721)
Qy 1 MetAlaAlaAspGlyTyrLeuProAspTyrLeuGluAspAsnLeuSerGluGlyIleArg 20
Db 2222 ATGGCTGCCGATGTTATCTCCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTCGC 2281
Qy 21 GluTyrTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnLysGlnAsp 40
Db 2282 GAGTGGTGGGACCTGAAACCTGGAGCCCCCGAAACCCAAAGCCCAACCCAGCAAGCAGGAC 2341
Qy 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp 60
Db 2342 AACGGCCGGGGTCTGGTGCTTCTGGCTACAGTACCTCGGACCCCTTCAACGGACTCGAC 2401
```

```
Qy 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80
Db 2402 AAGGGGAGCCCGTCAACGCGGGGACGCGCGCCCTCGAGCACGACAGGCTTACGAC 2461
Qy 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100
Db 2462 CAGCAGCTCAAAGCGGGTGACAATCCGTACCTGCGGTATAACCCAGCCGCGGAGTTT 2521
Qy 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120
Db 2522 CAGGAGCGTCTGCAAGAAGATACGTCAATTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG 2581
Qy 121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGluGlyAlaLysThrAlaPro 140
Db 2582 GCCAAGAAAGCGGGTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGCTCCT 2641
Qy 141 GlyLysLysArgProValGluGlnSerProGln--GluProAspSerSerSerGlyIle 159
Db 2642 GCAAAGAAGAGACCGGTAGAGCCGTCACTCAGCGTTCCTCCGACTCCTCCACGGGCATC 2701
Qy 160 GlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSer 179
Db 2702 GGCAAGAAAGGCCAGCAGCCCCGCCAGAAAGAGACTCAATTTCGGTTCAGACTGGCGACTCA 2761
Qy 180 GluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGly 199
Db 2762 GAGTCAGTCCCCCGACCTCAACCTCTCGGAGAACCTCCAGCAGCGCCCTTAGTGTGGGA 2821
Qy 200 ProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAsp 219
Db 2822 TCTGGTACAGTGGTTCAGGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGTGCCGAC 2881
Qy 220 GlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgVal 239
Db 2882 GGAGTGGGTAAATGCCCTCAGGAAATTTGGCATTGGCATTCACATGGTGGCGCAGAGTCA 2941
Qy 240 IleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGln 259
Db 2942 ATTACCACCCAGCACCCGAAACCTGGGCCCTGCCACCTACCAACCAACCCACCTCTACAAGCAA 3001
Qy 260 IleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrPro 279
Db 3002 ATCTCCAGTGAACACTGCAGGTAGTACCACGACGACACACCTACTTCGGCTACAGCACCCCC 3061
Qy 280 TrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTyrGlnArg 299
Db 3062 TGGGGGTATTTTGACTTTAACAGATTCCACTGCCACTTCTCACCACGTGACTGGCAGCGA 3121
Qy 300 LeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIle 319
Db 3122 CTCATCAACAACAACCTGGGGATTCCGGCCCCAAGAAAGCTGCGGTTCAGCTCTTCAACATC 3181
Qy 320 GlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSer 339
Db 3182 CAGGTCAAGGAGGTCACGACGAATGACGGCGTTACGACCATCGCTAATAACCTTACCAGC 3241
Qy 340 ThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHis 359
Db 3242 ACGATTCAAGTATTTCTCGGACTCGGAATACCAGCTGCGGTACGTCCTCGGCTCTGCGCAC 3301
Qy 360 GlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeu 379
Db 3302 CAGGGCTGCCTGCCCTCCGTTCCCGGGGACGCTCTTCATGATTCTCAGTACGGCTACCTG 3361
Qy 380 ThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhe 399
Db 3362 ACTCTCAACAATGGCAGTCAGTCTGTGGGACGTTCTCTCTTCTACTGCCTGGAGTACTTC 3421
Qy 400 ProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluVal 419
Db 3422 CCCTCTCAGATGCTGAGAACGGGCAACAACCTTTTGAGTTCAGCTACAGCTTCGAGGACGTG 3481
```



QY 420 PropHeHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIle 439  
Db 3482 CCTTCCACAGCAGCTACGCACACAGCCAGCGCTGGACCGGTGATGATCCCTCATC 3541  
QY 440 AspGlnTyrLeuTyrLeuAsnArgThrGln---AsnGlnSerGlySerAlaGlnAsn 458  
Db 3542 GACCAGTACTTGTACTACCTGGCCAGAACACAGAGTAACCCAGGAGGCACAGCTGGCAAT 3601  
QY 459 LysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrp 478  
Db 3602 CGGGAACCTGCAGTTTACCAGGGCGGGCCCTTCAACTATGGCCGAACAAGCCAAAGATTGG 3661  
QY 479 LeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 498  
Db 3662 TTACCTGGACCTTGCTTCCGGCAACAAAGAGTCTCCAAAACGCTGGATCAAAAACAACAAC 3721  
QY 499 SerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIle 518  
Db 3722 AGCAACTTTGCTTGGACTGGTGCCACCACCAATATCACCTGAACGGCAGAAAATCGTTGGTT 3781  
QY 519 AsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSer 538  
Db 3782 AATCCCGCGTCGCATGGCAACTCACAGGACGACGAGGCCGCTTTTCCCATCCAGC 3841  
QY 539 GlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsn---ThrAlaLeuAspAsn 557  
Db 3842 GGAGTCTGATTTTGGAAAA-----ACTGGAGCAACTAACAAAACATACATTGGAAAAT 3895  
QY 558 valMetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 577  
Db 3896 GTGTTAATGACAATGAAGAAGAAATTCGTCTACTAATCCTGTAGCCAGGAAGATAC 3955  
QY 578 GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis 597  
Db 3956 GGGATAGTCAGCAGCAACTTACAGCGGCTAATACTGCAGCCAGACACAAGTTGTCAAC 4015  
QY 598 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 617  
Db 4016 AACCAGGGAGCCTTACCTGGCATGGTCTGGCAGAACCGGGACGTGTACTGCAGGGTCCC 4075  
QY 618 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 637  
Db 4076 ATCTGGGCCAAGATTCTCACACGGATGGCAACTTTCACCGCTCTCCTTTGATGGCGGC 4135  
QY 638 PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn 657  
Db 4136 TTTGGACTTAAACATCCGCCTCCTCAGATCCTGATCAAGAACACTCCCGTTCCCGCTAAT 4195  
QY 658 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 677  
Db 4196 CCTCCGGAGGTGTACTCCTGCCAAGTTTGTCTGTTTCATCACACAGTACAGCACCCGA 4255  
QY 678 GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro 697  
Db 4256 CAAGTCAGCGTGGAAATCGAGTGGAGCTGCAGAAGGAAAAACAGCAAGCGCTGGAACCCG 4315  
QY 698 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 717  
Db 4316 GAGATTTCAGTACACCTCCAACCTTTGAAAAGCAGACTGGTGTGGACTTTGCCGTTGACAGC 4375  
QY 718 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736  
Db 4376 CAGGGTGTTTACTCTGAGCCTCGCCCTATTGGCACTCGTTACCTCACCCCGTAATCTG 4432

RESULT 14  
US-10-291-583-47  
; Sequence 47, Application US/10291583  
; Publication No. US20030138772A1  
; GENERAL INFORMATION:  
; APPLICANT: Gao, Guangping  
; APPLICANT: Wilson, James M.  
; APPLICANT: Alvira, Mauricio  
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus

; TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1el Sequences Identif  
; FILE REFERENCE: UPN-02735USA  
; CURRENT APPLICATION NUMBER: US/10/291,583  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: US 60/350,607  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/341,117  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: US 60/377,066  
; PRIOR FILING DATE: 2002-05-01  
; PRIOR APPLICATION NUMBER: US 60/386,675  
; PRIOR FILING DATE: 2002-06-05  
; NUMBER OF SEQ ID NOS: 120  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 47  
; LENGTH: 3128  
; TYPE: DNA  
; ORGANISM: 44.5  
US-10-291-583-47

Alignment Scores:  
Pred. No.: 0 Length: 3128  
Score: 3441.00 Matches: 626  
Percent Similarity: 91.06% Conservative: 46  
Best Local Similarity: 84.82% Mismatches: 64  
Query Match: 86.26% Indels: 2  
DB: 13 Gaps: 2

US-09-807-802A-13 (1-736) x US-10-291-583-47 (1-3128)

QY 1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsnLeuSerGluGlyIleArg 20  
Db 844 ATGCTGCCGATGGTTATCTCCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTCG 903  
QY 21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp 40  
Db 904 GAGTGGTGGGACTTGAAACCTGGAGCCCCGAAACCCAAAGCCAAACCCAGCAAAAGCAGGAC 963  
QY 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp 60  
Db 964 GACGCGGGGTCTGGTGTCTCTGGCTACAGTACCTCGGACCTTCAACGGACTCGAC 1023  
QY 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80  
Db 1024 AAGGGGAGCCCCGTCAACGGCGGACGACGCGGCCCTCGAGCAGCACCAAGGCCTACGAC 1083  
QY 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100  
Db 1084 CAGCAGCTCAAAGCGGTGACAATCCGTACCTGCGGTATAACCAACCGCCGAGCTTT 1143  
QY 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120  
Db 1144 CAGGACGCTCTCAAGAAGATACGTCTTTTGGGGCAACCTCGGGCGAGCAGTCTTCCAG 1203  
QY 121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro 140  
Db 1204 GCCAAGAAGCGGTTCTCGAACCTCTCGGTCTGGTTGAGNAAGCGCTAAGACGGCTCCT 1263  
QY 141 GlyLysLysArgProValGluGlnSerProGln---GluProAspSerSerSerGlyIle 159  
Db 1264 GGAAGAAGAGACCCGGTAGAGCATCACCCAGCGTTCTCCAGACTCTCTACGGGCATC 1323  
QY 160 GlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSer 179  
Db 1324 GGCAAGAAGGCCAGCAGCCCGGAAAAAAGAGACTCAACTTTGGGCAGACTGGCGACTCA 1383  
QY 180 GluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGly 199  
Db 1384 GAGTCAGTCCCGACCCCTCAACCAATCGGAGAACCCCCCGAGGCCCTCTGTGCTCGGA 1443  
QY 200 ProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnGluGlyAlaAsp 219  
Db 1444 TCTGTTACAATGGCTGCAGGGCGTGGCGCTCCAATGGCAGACAATAACGAAGGCCCGAC 1503



Db 844 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC 903

QY 21 GluTrpTrpAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp 40

Db 904 GAGTGGTGGGACTTGAACCTGGAGCCCCGAAACCCAAAGCCAAACGAGAAAAGCAGGAC 963

QY 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp 60

Db 964 GACGGCCGGGGTCTGGTCTCTTGGCTTACAAAGTACCTCGACCCCTTCAACGGACTCGAC 1023

QY 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80

Db 1024 AAGGGGAGCCCGTCAACGGCGGAGCGACGCGGCCCTCGAGCACGACAAGGCTACGAC 1083

QY 81 GlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsnHisAlaAspAlaGluPhe 100

Db 1084 CAGCAGCTCAAAGCGGGTGACAAATCCGTACCTGCGGTATAACCAACGCGGACGCGAGTTT 1143

QY 101 GlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeuGlyArgAlaValPheGln 120

Db 1144 CAGGAGCGTCTGCAAGAAAGATACGTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG 1203

QY 121 AlaLysLysArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaPro 140

Db 1204 GCCAAGAAGCGGGTTCTCGAACTCTCGGTCTGGTTGAGAAGGCGCTAAGACGGCTCCT 1263

QY 141 GlyLysLysArgProValGlnGlnSerProGln---GluProAspSerSerSerGlyIle 159

Db 1264 GGAAGAAGAGACCGGTAGAGCCATCACCCAGCGTTCTCCAGACTCTCTACGGGCATC 1323

QY 160 GlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSer 179

Db 1324 GGCAAGACAGCCAGCAGCCCCGGGAAAAGAGACTCAACTTTGGGCAGACTGGCGACTCA 1383

QY 180 GluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGly 199

Db 1384 GAGTCAGTCCCGACCCCTCAACCAATCGGAGAACCCCCCGAGGCCCTCTCTGGTCTGGGA 1443

QY 200 ProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAsp 219

Db 1444 TCTGGTACAATGGCTGCAGGCGGTGGCGCTCCAATGGCAGACAATAACGAAGGCGCCGAC 1503

QY 220 GlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgVal 239

Db 1504 GGAGTGGGTAGTTCTCAGGAAATTGGCATTTGCGATTCCACATGGCTGGCGGACAGAGTC 1563

QY 240 IleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGln 259

Db 1564 ATCACCACAGCACCCGAACCTGGGCCCTCCCCACCTACAAACACCACCTCTACAAGCAA 1623

QY 260 IleSerSerAlaSerThrGlyAlaSer---AsnAspAsnHisTyrPheGlyTyrSerThr 278

Db 1624 ATCTCCAACGGGACATCGGAGGAGGAGCAACCAACGACACACCTACTTCGGCTACAGCAC 1683

QY 279 ProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGln 298

Db 1684 CCCTGGGGGTATTTGACTTTAACAGATTCCACTGCCACTTCTCACACAGTACTGGCAG 1743

QY 299 ArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsn 318

Db 1744 CGACTCATCAACAACAACCTGGGGATTCCGGCCCAAGAGACTCAACTTCAAGCTCTTCAAC 1803

QY 319 IleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThr 338

Db 1804 ATCCAGGTCAAGAGGTCACGCAGAATGAAGGCACCAAGACCATCGCCAATAACCTTACC 1863

QY 339 SerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAla 358

Db 1864 AGCAGGATTCAGGTCTTTACGGACTCGGAATACCAAGCTCCCGTACGTCCTCGGCTCTGCG 1923

QY 359 HisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyr 378

Db 1924 CACCAGGGCTGCCCGCTCCGTTCCCGCGGAGCGTCTTCATGATTCTCAGTACGGGTAC 1983

QY 379 LeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyr 398

Db 1984 CTGACTCTGAACAACCGCAGTCAGGCCGTGGGCCGTTCTCTTCTACTGCCTGGAGTAC 2043

QY 399 PheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGlu 418

Db 2044 TTTCTCTTCTCAAAATGCGGAGAACGGGCAACAACCTTTGAGTTTCAGTACCAGTTTGAGGAC 2103

QY 419 ValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeu 438

Db 2104 GTGCCCTTTTCACAGCAGCTACGCGCATAGCCAAAGCCTGGACCGGCTGATGAACCCCTC 2163

QY 439 IleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsn 458

Db 2164 ATCAGCCAGTACCTGTACTACCTGTCTCGGACTCAGTCCACGGGAGGTACCGCAGGAAC 2223

QY 459 LysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrp 478

Db 2224 CAGCAGTTGCTATTTTCTCAGGCCGGGCTTAATAACATGTTCGGCTCAGGCCAAAACTGG 2283

QY 479 LeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsn 498

Db 2284 CTACCCGGGCCCTGTACTACGGCAGCAACCGCTCTCCACGACACTGTCCGCAAAATAACAAC 2343

QY 499 SerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIle 518

Db 2344 AGCAACTTTGCTTGGACCGGTGCCACCAAGTATCATCTGAATGGCAGAGACTCTCTGGTA 2403

QY 519 AsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSer 538

Db 2404 AATCCCGGTGTCGTATGGCAACGCAACAGGACGACGAAGCGGATTTTTCATCCAGC 2463

QY 539 GlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnVal 558

Db 2464 GGAGTCTTGATGTTGGGAAACAGGGAGCTGGAAAAAGACAACGTTGACTATAGCAGCGTT 2523

QY 559 MetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGly 578

Db 2524 ATGCTAACCACTGAGGAAGAATACTAAACCCCAACCCAGTGGCCACAGAACAGTACGGC 2583

QY 579 ThrValAlaValAsnPheGlnSerSerSerSerThrAspProAlaThrGlyAspValHisAla 598

Db 2584 GTGGTGGCCGATAACCTGCAACAGCAAAACGCCGCTCTTATTGTAGGGGCCGTCAACAGT 2643

QY 599 MetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIle 618

Db 2644 CAAGGAGCCTTACCTGGCATGGTCTGGCAGAACCGGACGTGTACCTGCAGGGTCTCTATC 2703

QY 619 TrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPhe 638

Db 2704 TGGGCCAAGATTCTCTCACAGGACGGCAACTTTCATCTCTTCGCGCTGATGGGAGGCTTT 2763

QY 639 GlyLeuLysAsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsnPro 658

Db 2764 GGACTGAAACACCCCGCTCTCTCAGATCCTGTATTAGAATAACACCTGTTCCCGCGGATCCT 2823

QY 659 ProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGln 678

Db 2824 CCAACTACCTTCAGTCAAGCCCAAGCTGGCGTCTGTTTCATCAGCAGTACAGCACCCGACAG 2883

QY 679 ValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGlu 698

Db 2884 GTCAGCGTGGAAATTGAATGGGAGCTGCAGAAAGAGAAAGACAGCAAGCGCTGGAACCCAGAG 2943

QY 699 ValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsn 718

Db 2944 ATTCAGTATACTTCCAATACTACTACAAATCTACAAATGTGAGACTTTGTGTCAATACTAG 3003

QY 719 GlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 736

Db 3004 GGTACTTATTCAGAGCCTCGCCCCCATTTGGCACCCCGTTACCTCACCCCGTAACCTG 3057



Sun Feb 15 18:36:55 2004

us-09-807-802a-13.rnpb

Page 21

Search completed: February 15, 2004, 02:04:02  
Job time : 852.807 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 14, 2004, 15:26:50 ; Search time 4084.43 Seconds  
(without alignments)  
4379.586 Million cell updates/sec

Title: US-09-807-802A-13  
Perfect score: 3989  
Sequence: 1 MAADGYLPDWLENDLSEGIR.....NNGLYTEPRPIGTRYLTRPL 736

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09807802/runat\_11022004\_175608\_15941/app\_query.fasta\_1.2389  
-DB=EST -QFMT=fastap -SUFFIX=rbt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09807802 @CGN 1 1 7257 @runat\_11022004\_175608\_15941 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID       | Description         |
|------------|-------|-------------|--------|----|----------|---------------------|
| 1          | 193.5 | 4.9         | 753    | 28 | BH115587 | BH115587 RPCI-24-3  |
| 2          | 147   | 3.7         | 2318   | 11 | BC014681 | BC014681 Homo sapi  |
| 3          | 139.5 | 3.5         | 803    | 29 | BZ265284 | BZ265284 CH230-375  |
| 4          | 134   | 3.4         | 856    | 28 | BH164736 | BH164736 ENTTTS28TF |
| 5          | 133   | 3.3         | 2146   | 11 | AK042727 | AK042727 Mus muscu  |
| 6          | 133   | 3.3         | 2295   | 11 | AK035953 | AK035953 Mus muscu  |
| 7          | 129   | 3.2         | 3943   | 11 | AK087510 | AK087510 Mus muscu  |
| c 8        | 125   | 3.1         | 640    | 14 | CA356072 | CA356072 628112 NC  |
| c 9        | 125   | 3.1         | 701    | 14 | CA363789 | CA363789 638563 NC  |
| c 10       | 125   | 3.1         | 703    | 14 | CA378088 | CA378088 656813 NC  |
| c 11       | 125   | 3.1         | 745    | 12 | BJ139005 | BJ139005 BJ139005   |
| 12         | 125   | 3.1         | 754    | 10 | BG207226 | BG207226 RST26694   |
| 13         | 125   | 3.1         | 2019   | 11 | AK014322 | AK014322 Mus muscu  |
| 14         | 124   | 3.1         | 925    | 13 | BQ440075 | BQ440075 AGENCOURT  |
| 15         | 123.5 | 3.1         | 645    | 12 | BI911451 | BI911451 603063370  |
| 16         | 123.5 | 3.1         | 3512   | 11 | AK038988 | AK038988 Mus muscu  |
| 17         | 123.5 | 3.1         | 4763   | 11 | AK083225 | AK083225 Mus muscu  |
| c 18       | 122.5 | 3.1         | 756    | 28 | AQ742711 | AQ742711 HS-5386 B  |
| 19         | 121.5 | 3.0         | 925    | 28 | BH152154 | BH152154 ENTPV12TR  |
| 20         | 121.5 | 3.0         | 1418   | 13 | BQ686477 | BQ686477 AGENCOURT  |
| 21         | 121.5 | 3.0         | 3689   | 11 | AK031988 | AK031988 Mus muscu  |
| 22         | 120.5 | 3.0         | 567    | 14 | CB190643 | CB190643 pl29c03.Y  |
| c 23       | 120.5 | 3.0         | 629    | 14 | CA408136 | CA408136 pk09b07.x  |
| c 24       | 120.5 | 3.0         | 902    | 29 | CNS03388 | AL225809 Tetraodon  |
| 25         | 120.5 | 3.0         | 2080   | 11 | BC001541 | BC001541 Homo sapi  |
| c 26       | 120   | 3.0         | 812    | 28 | BH600978 | BH600978 BOGUZ82TF  |
| 27         | 120   | 3.0         | 937    | 12 | BM802157 | BM802157 AGENCOURT  |
| 28         | 120   | 3.0         | 4782   | 11 | AK048546 | AK048546 Mus muscu  |
| 29         | 119.5 | 3.0         | 3329   | 11 | AK033012 | AK033012 Mus muscu  |
| 30         | 119   | 3.0         | 919    | 13 | BX382214 | BX382214 BX382214   |
| 31         | 118.5 | 3.0         | 973    | 29 | AG072149 | AG072149 Pan trogl  |
| 32         | 118.5 | 3.0         | 1035   | 12 | BQ061577 | BQ061577 AGENCOURT  |
| 33         | 118   | 3.0         | 3344   | 11 | BC032608 | BC032608 Homo sapi  |
| 34         | 117   | 2.9         | 5809   | 11 | AK076994 | AK076994 Mus muscu  |
| 35         | 116.5 | 2.9         | 1085   | 13 | BQ645964 | BQ645964 AGENCOURT  |
| 36         | 116.5 | 2.9         | 2249   | 11 | AK049740 | AK049740 Mus muscu  |
| 37         | 116   | 2.9         | 1604   | 11 | BC030211 | BC030211 Homo sapi  |
| 38         | 115.5 | 2.9         | 1042   | 13 | BQ642977 | BQ642977 AGENCOURT  |
| 39         | 115.5 | 2.9         | 1065   | 13 | BQ073024 | BQ073024 AGENCOURT  |
| c 40       | 115   | 2.9         | 1133   | 13 | BQ716249 | BQ716249 AGENCOURT  |
| 41         | 115   | 2.9         | 2931   | 11 | AK019448 | AK019448 Mus muscu  |
| 42         | 115   | 2.9         | 4784   | 11 | AK041115 | AK041115 Mus muscu  |
| 43         | 114.5 | 2.9         | 641    | 10 | BF984198 | BF984198 602307620  |
| c 44       | 114.5 | 2.9         | 936    | 12 | BI838324 | BI838324 603083262  |
| 45         | 114.5 | 2.9         | 1023   | 13 | BU850074 | BU850074 AGENCOURT  |

ALIGNMENTS

RESULT 1  
BH115587  
LOCUS BH115587 753 bp DNA linear GSS 19-JUL-2001  
DEFINITION RPCI-24-358F16.TV RPCI-24 Mus musculus genomic clone RPCI-24-358F16  
, genomic survey sequence.  
ACCESSION BH115587  
VERSION BH115587.1 GI:14954954  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 753)

AUTHORS Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,  
Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,  
Russell,D., de Jong,P. and Fraser,C.M.  
TITLE Mouse BAC End Sequences from Library RPCI-24  
JOURNAL Unpublished  
COMMENT Other CSSs: RPCI-24-358F16.TJ  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org  
Clones are derived from the mouse BAC library RPCI-24. For BAC  
library availability, please contact Pieter de Jong  
(pdejong@mail.cho.org). Clones may be purchased from BACPAC  
Resources (http://www.choi.org/bacpac/orderingframe.htm). BAC end  
page: http://www.tigr.org/tdb/bac\_ends/mouse/bac\_end\_intro.html  
Plate: 358 row: F column: 16  
Seq primer: T7  
Class: BAC ends.

Location/Qualifiers  
1. .753  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCI-24-358F16"  
/sex="Male"  
/cell\_type="Spleen/Brain"  
/clone\_lib="RPCI-24"  
/note="Vector: pTARBAC1; Site 1: BamH1; Site 2: BamH1;  
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The  
library was cloned in the pTARBAC1 cloning vector at the  
BamH1 sites using MboI partially digested male C57BL/6J  
DNA."

BASE COUNT 211 a 178 c 167 g 197 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 9.71e-07 Length: 753  
Score: 193.50 Matches: 68  
Percent Similarity: 46.74% Conservative: 18  
Best Local Similarity: 36.96% Mismatches: 64  
Query Match: 4.85% Indels: 36  
DB: 28 Gaps: 5

US-09-807-802A-13 (1-736) x BH115587 (1-753)

QY 91 LeuArgTyrAsnHisAlaAspAlaGluPheGlnGlu-ArgLeuGlnGluAspThrSerPh 110  
Db 212 CTCCATATAAATCACACAGAGCAGAGTTTCAGGAGAAACTCCAAGTACAATCTTTT 271  
QY 110 eGlyGlyAsnLeuGlyArgAlaValPheGlnAlaLysLysArgValLeuGluProLeuG1 130  
Db 272 T-----TTGGCAGGAACCGTGCCAAATGCCAAAGAAAGGGTTCTCAAAACCCATCGG 322  
QY 130 YLeuValGluGluGlyAlaLysThrAlaProGlyLysLys-----ArgProValGluG1 148  
Db 323 CTTGCTCTAAGAGGATTTCAGCATGGCTCTCTGGGAAGAAATAGATTTCGACCTTA-AACTC 381  
QY 148 nSerPro-----GlnGluProAspSerSerSerGlyI1 159  
Db 382 CTCCCCCACATGAAGGAAGTACACTCGCTCTTTTCAGAGTTATGCCAAAACAGTAGAAG 441  
QY 159 eGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspse 179  
Db 442 TGGAGAGGCTGGCTCATATTAGTCTATGACACAGATGTAGCAGTTTCCAGCAGATATCCTC 501  
QY 179 rGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValG1 199  
Db 502 CAATTTA-----508

QY 199 yProThrThrMetAlaSerGly---GlyGlyAlaProMetAlaAspAsnAsnGluGlyAl 218  
Db 509 -GACACTTCTATAATCTCTGGAGCTGGAGGTCTACTAATGGCAACAACACACAGAC-AC 566  
QY 218 aAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspAr 238  
Db 567 TGATGGAGTGAGCAATGCCATGGATGATGGCATTCATCCCAAGTGTATGGTAGACTG 626  
QY 238 gValIleThrThrSerThrArgThrTrpAlaLeuProThrTyxAsnAsnHisLeuTyxLy 258  
Db 627 AGTCATTACCTAATCTCCCGAGACCTGGTGCTCTGCCCAACTCCAACATAACATTTCACAA 686  
QY 258 sGlnIleSer 261  
Db 687 CTATATGAAC 696

RESULT 2  
BC014681  
LOCUS BC014681.1 GI:15928395 2318 bp mRNA linear HTC 04-OCT-2001  
DEFINITION Homo sapiens, Similar to osa, clone IMAGE:3866722, mRNA.  
ACCESSION BC014681  
VERSION BC014681.1  
KEYWORDS HTC.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2318)  
AUTHORS Strausberg,R.  
TITLE Direct Submission  
JOURNAL Submitted (01-OCT-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov  
COMMENT Contact: MGC help desk  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: http://www.hgsc.bcm.tmc.edu/cdna/  
Contact: villalona@bcm.tmc.edu.  
Villalona, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,  
A.M., Holloway, M., Telford, B, Hodgson, A., Bouck, J., Yu, W.,  
Muzny,D.M., Gibbs,R.A.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Series: IRAK Plate: 21 Row: 1 Column: 9  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Genomescan gene prediction  
This clone has the following problem: frame shifted.

FEATURES  
source  
1. .2318  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3866722"  
/tissue\_type="Eye, retinoblastoma"  
/clone\_lib="NIH MGC\_67"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"

BASE COUNT 637 a 680 c 544 g 457 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.0559 Length: 2318  
Score: 147.00 Matches: 172  
Percent Similarity: 32.27% Conservative: 91  
Best Local Similarity: 21.10% Mismatches: 273



|   |       |  |      |
|---|-------|--|------|
| Query Match:                                  | 3.69% | Indels:  | 281  |
| DB:   | 11    | Gaps:  | 45   |
| US-09-807-802A-13 (1-736) x BC014681 (1-2318) |       |  |      |
| QY  | 1     | MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsn-----LeuSerGluGly  | 18   |
| Db  | 388   | ATGGCAGATCCAGGAATGATGAGTCTTTTGGCGAGGATGGGAATATTTTCAGTGAAGGT  | 447  |
| QY  | 19    | IleArgGluTrpTrpAspLeuLysProGlyAlaProLys-----ProLysAla        | 34   |
| Db  | 448   | CTTGAAGGCCTCGAGAAATGT-----GGTTACCCCGGAAATCCAGTAAATCCTATGGGT  | 501  |
| QY  | 35    | AsnGlnGlnLysGlnAspAspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGly | 54   |
| Db  | 502   | CAGCAAAATGCCAATAGACCAAGGCTTTGGCTCTTTTACAGCCATCCCTTCATCAT     | 555  |
| QY  | 55    | ProPheAsnGlyLeuAspLysGlyProValAsnAlaAlaAspAlaAlaLeuGlu       | 74   |
| Db  | 556   | -----CCTTCAACTAATCAAAAATCAACAAAGCTGACACATTTTGAT              | 597  |
| QY  | 75    | HisAspLysAlaTyrAspGlnGln---LeuLysAlaGlyAspAsnProTyrLeuArgTyr | 93   |
| Db  | 598   | CACATAATCAGTATGAACAACAAGATGCTGTGATCGATCAGCCG-----            | 645  |
| QY  | 94    | AsnHisAlaAspAlaGluPheGlnGluArgLeuGlnGluAspThrPheGlyGlyAsn    | 113  |
| Db  | 646   | -----AACAGATGATGAGCAACACCCCT-----GGGAAC                      | 675  |
| QY  | 114   | LeuGlyArgAlaValPheGlnAlaLysLysArgValLeuGluProLeuGlyLeuValGlu | 133  |
| Db  | 676   | -----GGACTC-----   | 681  |
| QY  | 134   | GluGlyAlaLysThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluPro | 153  |
| Db  | 682   | -----GCGTCTCCGCACCTCGCAGTATCACACCCCTCCCGTTCTCAGGTGCC         | 729  |
| QY  | 154   | AspSerSerGlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPhe    | 173  |
| Db  | 730   | CATGGTGGCAGTGGTGGCGGTGAGTGGGTGTCTACCTTGGCATGCAGAATGAGAGGCAT  | 789  |
| QY  | 174   | GlyGlnThr---GlyAspSerGluSer-----ValProAsp                    | 184  |
| Db  | 790   | GGCAATCCTTTGTGGACAGCAGCTCCATGTGGGCCCCCAGGGCTGTTTCAGGTACCAGAC | 849  |
| QY  | 185   | -----ProGlnProLeuGlyGluProProAlaThrProAla                    | 196  |
| Db  | 850   | CAGATACGAGCCCCCTACCAGCAGCAGCCACAGCCGAGCCGACCCG---CAGCCGGCT   | 906  |
| QY  | 197   | AlaValGlyPro-----ThrThrMet                                   | 203  |
| Db  | 907   | CCGTCGGGGCCCCCTGCACAGGGCCACCCTCAGCACATGCAGCAGATGGGCAGCTATATG | 966  |
| QY  | 204   | AlaSerGlyGly-----GlyAlaPro-----MetAlaAspAsn                  | 214  |
| Db  | 967   | GCACGTGGGGATTTTTCATGCAGCAGCATGGTTCAGCCACAGCAGAGGATGAGCCAGTTT | 1026 |
| QY  | 215   | AsnGluGlyAlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrp | 234  |
| Db  | 1027  | TCCCAAGGCCAAGAGGGCCTCAATCAG-----                             | 1053 |
| QY  | 235   | LeuGlyAspArgValIleThrThrSerThrArgThrTroAlaLeuProThrTyrAsnAsn | 254  |
| Db  | 1054  | ---GGAATCCTTTTATTGCCACCTCAGGA-----CCTGGCCACTGTGTC            | 1095 |
| QY  | 255   | HisLeuTyrLysGlnIleSerSerAlaSerThrGlyAlaSerAsnAsp-----        | 270  |
| Db  | 1096  | CACGTGCCCCCAGCAGAGTCCAGCATGGCACCTTCTTGGTCACTCGGTGCACAGTTTC   | 1155 |
| QY  | 271   | AsnHisTyrPheGlyTyrSerThrProTrpGlyTyrPheAspPheAsnArgPheHisCys | 290  |
| Db  | 1156  | CATCAACCAACC-CTCTACTGTCTCTCCATGG-----AGAATCCGTTC             | 1196 |
| QY  | 291   | -His-----PheSerProArgAspTrpGlnArgLeuIleAsnAsnAsnTrpGlyPh     | 307  |
| Db  | 1197  | CCACAGTCCCAGATTCTCCCGAATCCTCCCCAACAA-----GGGGCTGT            | 1241 |
| QY  | 307   | eArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLysGluValThrThrAs | 327  |
| Db  | 1242  | TAGGCCGCAAAACCTTAACCTTAGTTCTCGGAGCCAGACAGTCCCCTCTCTACTATAAA  | 1301 |
| QY  | 327   | nAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGlnValPheSerAspSe | 347  |
| Db  | 1302  | CAAC-----TCAGGGCA  | 1313 |
| QY  | 347   | rGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCysLeu-----       | 363  |
| Db  | 1314  | GTATTCTCGATATCCTTAC-----AGTAACCTAAATCAGGGATTAGTTAACAAATACAGG | 1367 |
| QY  | 364   | -----ProProPheProAlaAspValPheMetI                            | 373  |
| Db  | 1368  | GATGAATCAAAATTTAGGCCTTACAAATAATACTCCATGAATCAGTCCGTA-----     | 1419 |
| QY  | 373   | eProGlnTyrGlyTyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPh | 393  |
| Db  | 1420  | -CCAAGATAC-----CCCAATGCTGTAGGA-----                          | 1443 |
| QY  | 393   | eTyrCysLeuGluTyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSe | 413  |
| Db  | 1444  | -----TTCCCATCA-----AACAGTGGTCAAGGACTA-----                   | 1470 |
| QY  | 413   | rTyrThrPheGluGluValPhePheHisSerSerTyrAlaHisSerGlnSerLeuAspAr | 433  |
| Db  | 1471  | -----ATGCACCAGAGCCCATCCACCCAGTGGCTCACTTAACCAATGAACACACA      | 1523 |
| QY  | 433   | gLeuMetAsnProLeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSe | 453  |
| Db  | 1524  | AACATATGCATCCT-----TCACAGCCTCAGGGAACCTTA                     | 1556 |
| QY  | 453   | rGlySerAlaGlnAsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerVa | 473  |
| Db  | 1557  | TGCCTCTCCACCTCCCATGTCCCATGAAAGCAATAGTAATCCAGCAGGC-----AC     | 1610 |
| QY  | 473   | IglnProLysAsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLy | 493  |
| Db  | 1611  | TCCTCTCCACAAGTCAGGCCCGGGAAG-TGCTGG-----                      | 1643 |
| QY  | 493   | sThrAspAsnAsnAsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnG1 | 513  |
| Db  | 1644  | -----GATACCAATGGAAGTTGGCAG---TTATCCAAATAT-----GC             | 1678 |
| QY  | 513   | yArgGluSerIleIleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLy | 533  |
| Db  | 1679  | CCCATCTCTCAGCCATCTCACCAGCCCTGGTGGCCATGGGAATCGGACAGAGGAA----- | 1733 |
| QY  | 533   | sPhePheProMetSerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnTh | 553  |
| Db  | 1734  | -----TATGGGCCCCAGAAACATGCAGCAGTCTCGTC                        | 1765 |
| QY  | 553   | rAlaLeuAspAsnValMetIleThrAsp-----GluGluGluIleLys-----        | 567  |
| Db  | 1766  | CATTTATAGGCATGTCTCGGCACCAAGGAATTGACTGGGCACATGAGGCCAATGGTT    | 1825 |
| QY  | 568   | -----AlaThrAsnProValAlaThrGluArgPheGlyTh                     | 579  |
| Db  | 1826  | GTCTGTGTGTGGCCTTGGAGACCCCAAGCAATCCAGGAACGACTGATACCTGGCCAAC   | 1885 |
| QY  | 579   | rValAlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMe    | 599  |
| Db  | 1886  | AACATCC-----TGGTCAACAGCCATCTTTTCAGCAGTTGCCAACCTGTCTCCAC      | 1936 |
| QY  | 599   | tGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTr | 619  |
| Db  | 1937  | TGCAGCCTCACCCGGG-----CTTGACACCACCATCTT                       | 1969 |
| QY  | 619   | pAlaLysIle-ProHisThrAsp-----GlyHisPheHisProSerProLeuMetGlyG  | 637  |

Db 1970 CACCTCCACACCCCTCATCACAGCCTTGGGCACAGCTCCACCCATCACCC----- 2019

Qy 637 lyPheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaA 657

Db 2020 -----CAGAACACCCCGCAGAAAGTG-----CCTGTGCATCAGC 2053

Qy 657 snProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrG 677

Db 2054 ATTCCCGTCGGAGCCCTTCTAGAGAAACCCAGTGC CGGATATGACTCAGGTTAGTGGAC 2113

Qy 677 lyGlnValSerVal-----GluIleGluTyrGluLeuGlnL 689

Db 2114 CGAATGCTCAGCTAGTGAAGAGTGATGATTACCTGCCATCAATAGAACAGCCACAAC 2173

Qy 689 ysGluAsnSerLysArgTyrAsnProGluValGlnTyrThrSerAsnTyrAlaLysSera 709

Db 2174 AAAAGAAGAAGAAAGAAACCAACCACATTGTAGCAGAGGATCCCAGTAAAGTTTGTG 2233

Qy 709 laAsnValAspPheThr-----ValAspAsnAsnGlyLeu 720

Db 2234 GTAAAGATGACTTCCTGTGGGTAGATAACCAAGAACTA 2274

RESULT 3

BZ265284 803 bp DNA linear GSS 15-OCT-2002

LOCUS CH230-375A22.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone

DEFINITION CH230-375A22, genomic survey sequence.

ACCESSION BZ265284

VERSION BZ265284.1 GI:23978528

KEYWORDS GSS.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 803)

AUTHORS Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn ,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de Jong,P. and Fraser,C.M.

TITLE Rat BAC End Sequences from Library CHORI-230 MboI segment

JOURNAL Unpublished

COMMENT Other GSSs: CH230-375A22.TV

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230

(<http://www.chori.org/bacpac/rat230.htm>). For BAC library availability, please contact Pieter de Jong ([pdejong@mail.cho.org](mailto:pdejong@mail.cho.org)). Clones may be purchased from BACPAC Resources

([http://www.chori.org/bacpac/orering\\_information.htm](http://www.chori.org/bacpac/orering_information.htm)). BAC end page: [http://www.tigr.org/tdb/bac\\_ends/rat/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html)

Plate: 375 row: A column: 22

Seq primer: SP6

Class: BAC ends.

FEATURES

source

1. .803

/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/strain="BN/SnHsd/MCW"

/db\_xref="taxon:10116"

/clone="CH230-375A22"

/sex="Female"

/cell\_type="Brain"

/clone\_lib="CHORI-230 Segment 2"

/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI; CHORI-230 Rat (BN/SnHsd/MCW) BAC library produced by Pieter de Jong"

BASE COUNT 249 a 177 c 178 g 199 t

ORIGIN

Alignment Scores: 0.0593 Length: 803

Pred. No.: 139.50 Matches: 68

Score: 42.98% Conservative: 33

Percent Similarity: 28.94% Mismatches: 76

Best Local Similarity: 3.50% Indels: 60

Query Match: 29 Gaps: 10

DB:

US-09-807-802A-13 (1-736) x BZ265284 (1-803)

Qy 3 AlaAspGlyTyrLeuProAspTyrLeu-----GluAspAsnLeuSerGluGlyIleArg 20

Db 72 TCTGAGGGAACCTGCGCAAGACTGGTTGGTAGCAAGTCTAACCTGAGC-----TTAAAG 125

Qy 21 GluTyrTyrAspLeuLysProGlyAlaProLysProLysAlaAsnGlnGlnLysGlnAsp 40

Db 126 CAG-----GATAGCAAAAACACTACAACTAACCCCAATTCTAATCAGGAAACCAAGTC 179

Qy 41 AspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGlyProPheAsnGlyLeuAsp 60

Db 180 AATATATCTGCTCCTGCTTAGGTGC-TATAAGTACTTAGGGCTGCCAAAC----- 229

Qy 61 LysGlyGluProValAsnAlaAlaAspAlaAlaLeuGluHisAspLysAlaTyrAsp 80

Db 230 -----CAGTGATTTAATTCAGCAGAC----- 250

Qy 81 GlnGlnLeuLysAlaGlyAspAsnPro-----TyrLeuArgTyr 93

Db 251 ---ATGGTGAAGATCAAGACACACCATCACTCCTCAAGCATAGACTCATATCAAGTAC 307

Qy 94 AsnHisAlaAspAlaGluPheGlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsn 113

Db 308 AGCCTGTGCAAGCAGAAATTCAAGAGAAAAATCCAAACAGATGCTTTTCCAGAGGCAAT 367

Qy 114 LeuGlyArgAlaValPheGlnAlaLysLysArgValLeuGluProLeuGlyLeuValGlu 133

Db 368 CTCAACCAAGTACTCTCCAGGCTAAAGAGAGACTCTTGATCCCTTTGACCTGGTTAA- 426

Qy 134 GluGlyAlaLysThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluPro 153

Db 427 AAAGGTTGT-----CCA 438

Qy 154 AspSerSerGlyIleGlyLysThrGlyGlnGlnProAlaLysLys-----ArgLeu 171

Db 439 GAGATGCTCTCTGGCAGCCAGCAATGTACTCTTTTACCTTCAGCAAAGCTTCACAGAACT 498

Qy 172 AsnPheGlyGlnThrGlyAspSerGluSerValProAspProGlnProLeuGlyGlu--- 190

Db 499 AATCACTGAAGATGATAAGGTCACCTGAACATATATTAACCAACCTGGATCCCGTCATCT 558

Qy 191 -----ProProAlaThrProAlaAlaValGlyProThrThrMet 203

Db 559 GGTATGGTGTTCACAATAAGCAACTGCAAAACCCAGCAAACTTGGGGTCTCTTACAATG 618

Qy 204 Ala-----SerGlyGlyAlaProMetAlaAspAsnAsnGlu 216

Db 619 TCTCTTACACCAGAGAGGTGTTCACACAAATGACAATAACCAAG 663

RESULT 4

BH164736 856 bp DNA linear GSS 24-SEP-2001

LOCUS ENTTS28TF Entamoeba histolytica Sheared DNA Entamoeba histolytica

DEFINITION genomic, genomic survey sequence.

ACCESSION BH164736

VERSION BH164736.1 GI:15738174

KEYWORDS GSS.

SOURCE Entamoeba histolytica

ORGANISM Entamoeba histolytica

REFERENCE Eukaryota; Entamoebidae; Entamoeba.

AUTHORS Loftus,B., Wang,Z., Van Aken,S. and Fraser,C.

TITLE Determination of clone end sequences from Entamoeba histolytica  
HMI:IMSS sheared DNA library (2001)  
JOURNAL Unpublished  
COMMENT Contact: Brendan J Loftus  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 3543  
Email: bjloftus@tigr.org  
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared DNA library  
Seq primer: M13-Forward  
Class: shotgun  
High quality sequence start: 17  
High quality sequence stop: 629.  
Location/Qualifiers  
..1..856  
/organism="Entamoeba histolytica"  
/mol\_type="genomic DNA"  
/strain="HMI:IMSS"  
/db\_xref="taxon:5759"  
/clone\_lib="Entamoeba histolytica Sheared DNA"  
/note="Vector: pHO51; Site\_1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barell, Oxford University Press, 1999)."

FEATURES  
source  
..1..856

BASE COUNT 347 a 229 c 89 g 191 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.197 Length: 856  
Score: 134.00 Matches: 88  
Percent Similarity: 33.42% Conservative: 44  
Best Local Similarity: 22.28% Mismatches: 140  
Query Match: 3.36% Indels: 123  
DB: 28 Gaps: 19  
US-09-807-802A-13 (1-736) x BH164736 (1-856)

QY 297 TrpGlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPhelysLeu 316  
|||::: ||||| |||||  
Db 9 TGGAGACACTCATTTGGTAAAT-----TCTTCATTGCTGGA 47  
QY 317 PheAsnIleGlnValLysGluValThrThrAsnAspGlyValThr----- 332  
||| ||||| ||||| |||||  
Db 48 TTTAATACTAACCAACGTCGGGAACAATAATCCGTTCAATTCTACTGGAAGTTCAGTC 107  
QY 333 ---IleAlaAsnAsnLeuThrSerThrValGlnValPheSerAspSerGluTyrGlnLeu 351  
||| ||||| ||||| |||||  
Db 108 TCAGGAGCAATAACCCCTTTGCTACTACTCAA-----AACAAATACAACAAC 155  
QY 352 ProTyrValLeuGlySerAlaHisGlnGlyCysLeuProPheProAlaAspValPhe 371  
|||::: ||||| ||||| |||||  
Db 156 CCTTTCACAACAGGAACAATAAC-----CCATTAAAT----- 188  
QY 372 MetIleProGlnTyrGlyTyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSer 391  
||| ||||| ||||| |||||  
Db 189 -----ACTACTAACATACTACTAACTCA----- 212  
QY 392 SerPheTyrCysLeuGluTyrPheProSerGlnMetLeuArgThrGlyAsnAsnPhethr 411  
||| ||||| ||||| |||||  
Db 213 -----ACAACTAATCCATTAAAT 230

QY 412 PheSerTyrThrPheGluGluValProPheHisSerSerTyrAlaHisSerGlnSerLeu 431  
::: ||||| ||||| |||||  
Db 231 ACTAATACTACAAGCACCACCAATCCATTAAATACAATACTACTAAT----- 278  
QY 432 AspArgLeuMetAsnProLeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsn 451  
||| ||||| ||||| |||||  
Db 279 -----AATCCATTT-----AGTACTAATAAC 299  
QY 452 GlnSerGlySerAlaGlnAsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMet 471  
::: ||||| ||||| |||||  
Db 300 ACTACCAATGGAACAACAAT-----TTATTAAATCAA---ACTCCTTCTTCTAAT 347  
QY 472 SerValGlnProLysAsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSerLys 491  
::: ||||| ||||| |||||  
Db 348 ACTATAACACCAGGAAAT-----AATACAACACTACT 377  
QY 492 ThrLysThrAspAsnAsnAsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeu 511  
||| ||||| ||||| |||||  
Db 378 ACAACAGGAAGTAATCCATTGTTGTAATTTCACTACCACATAAT----- 419  
QY 512 AsnGlyArgGluSerIleIleAsnProGlyThrAlaMetAlaSerHisLysAspAspGlu 531  
||| ||||| ||||| |||||  
Db 420 AATACTACCTCCAGTACAACCTTCAACAGGAACAACCATCAACAACAGGAAGTAATCCATTT 479  
QY 532 AspLysPhePheProMetSerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSer 551  
||| ||||| ||||| |||||  
Db 480 GGTAATTTCACTCCCTCAAACCTCAGCTCCAAACAACAGGAATAATACAACAGGAACAGCT 539  
QY 552 AsnThrAlaLeuAspAsnValMetIleThrAspGluGluIleLysAlaThrAsnPro 571  
||| ||||| ||||| |||||  
Db 540 ACAACCACAGGAATAATAACAACCATCACA-----ACAGGAAGTAATCCA 584  
QY 572 ValAlaThrGluArgPheGlyThrValAlaValAsnPheGlnSerSerSerThrAspPro 591  
||| ||||| ||||| |||||  
Db 585 -----TTTGGT-----AATTTCACTCCCTCAAACAACAGGAACACTCCC 617  
QY 592 AlaThrGlyAspValHisAlaMetGlyAlaLeuProGlyMetValTyrGlnAspArgAsp 611  
||| ||||| ||||| |||||  
Db 618 ACCACAGGGAAT-----AATATCACAGGAACAACACTACCACAACAGGAAT 662  
QY 612 ValTyrLeuGlnGlyProIleTrpAlaLysIleProHisThrAspGlyHisPheHisPro 631  
||| ||||| ||||| |||||  
Db 663 AAATACACAGGA-----ACAGGAAGATATCCATTT-----GGTAATTTCACTACT 707  
QY 632 SerProLeuMetGlyGlyPheGlyLeuLysAsnProProGlnIleLeuIle-LysAs 651  
||| ||||| ||||| |||||  
Db 708 TCAAAACTCACTACCCACACAGGAACAACAACACTACACAGGAATAACACACTACAGGA 767  
QY 651 nThrProValProAlaAsnProProAlaGluPheSerAlaThrLysPheAlaSerPheI 671  
||| ||||| ||||| |||||  
Db 768 CAACCCCAACACCGGGAACACCCCAACACCGGAACACCCCAAC----- 810  
QY 671 eThrGlnTyrSerThrGlyGlnValSerValGluIleGluTyr 685  
||| ||||| ||||| |||||  
Db 811 -----ACAGGAAGGATCCTTTTGGAAATTTTCATAC 840

RESULT 5  
AK042727  
LOCUS AK042727 2146 bp mRNA linear HTC 05-DEC-2002  
DEFINITION Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:A730019I05 product:hypothetical protein, full insert sequence.  
ACCESSION AK042727  
VERSION AK042727.1 GI:26335284  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE 1  
AUTHORS Carninci, P. and Hayashizaki, Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)







VERSION AK035953.1 GI:26084922  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1  
AUTHORS Carninci,P. and Hayashizaki,Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PUBMED 10349636  
REFERENCE 2  
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE 20499374  
PUBMED 11042159  
REFERENCE 3  
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
MEDLINE 20530913  
PUBMED 11076861  
REFERENCE 4  
AUTHORS Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seva,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S. and Hayashizaki,Y.  
TITLE Functional annotation of a full-length mouse cDNA collection  
JOURNAL Nature 409 (6821), 685-690 (2001)  
MEDLINE 21085660  
PUBMED 11217851  
REFERENCE 5  
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
JOURNAL Nature 420, 563-573 (2002)  
REFERENCE 6  
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,

Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,P., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.  
TITLE Direct Submission  
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)  
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL:http://genome.gsc.riken.go.jp/  
URL:http://fantom.gsc.riken.go.jp/  
FEATURES  
source Location/Qualifiers  
1..2295  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="CS7BL/6J"  
/db\_xref="FANTOM\_DB:9630020J18"  
/db\_xref="taxon:10090"  
/clone="9630020J18"  
/tissue type="cerebellum"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev stage="16 days neonate"  
misc\_feature 1..2295  
/note="hypothetical protein (evidence: truncated-Longest-ORF)"  
BASE COUNT 571 a 743 C 586 G 395 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.938 Length: 2295  
Score: 133.00 Matches: 162  
Percent Similarity: 31.04% Conservative: 82  
Best Local Similarity: 20.61% Mismatches: 270  
Query Match: 3.33% Indels: 276  
DB: 11 Gaps: 40  
US-09-807-802A-13 (1-736) x AK035953 (1-2295)  
Qy 1 MetAlaAlaAspGlyTyrLeuProAspTrpLeuGluAspAsn-----LeuSerGluGly 18  
Db 402 ATGGCGGACCCAGGATGATGAGTCTTTTGGCAGGATGGAGTCTGTTCAGCGAAGGC 461  
Qy 19 IleArgGluTrpTrpAspLeuLysProGlyAlaProLys-----ProLysAla 34  
Db 462 CTCGAAGGCCTCGGGGAATGT-----GGCTACCCGAAACCCCGTGAACCCCATGGGT 515  
Qy 35 AsnGlnGlnLysGlnAspAspGlyArgGlyLeuValLeuProGlyTyrLysTyrLeuGly 54  
Db 516 CAGCAGATGCCCATAGACCAAGGTTTTCCTTCCTTACAGCCATCCTCCACCATCTTCA 575  
Qy 55 PropheAsnGlyLeuAspLysGlyGluProValAsnAlaAlaAspAlaAlaLeuGlu 74  
Db 576 CCTAATCAG-----AATCAAACCAAGTTGACCCATTGTGAT 611  
Qy 75 HisAspLysAlaTyrAspGlnGlnLeuLysAlaGlyAspAsnProTyrLeuArgTyrAsn 94  
Db 612 CACTATAGTCAGTATGAACAGAGATGCATCTGATGGATCAGCCT----- 656  
Qy 95 HisAlaAspAlaGluPheGlnGluArgLeuGlnGluAspThrSerPheGlyGlyAsnLeu 114  
Db 657 -----AACCGAATGATGGGCAGCGCCCC 680  
Qy 115 GlyArgAlaValPheGlnAlaLysLysArgValLeuGluProLeuGlyLeuValGluGlu 134  
Db 681 GCGAACGGATTGGCGTCTCCGCACTCACAGTATCACACCCCT----- 722



QY 135 GlyAlaLysThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAsp 154  
Db 723 -----  
QY 155 SerSerSerGlyIleGlyLysThrGlyGlnGlnProAlaLysLys----- 169  
Db 753 GGTGGCGGTcAGATGGGAGT-CTATCCCGGCATCCAGAACGAAAGGCACGGGCAATCTTT 811  
QY 170 -----ArgLeuAsnPheGlyGlnThrGlyAspSerGluSerValProAspProGln 186  
Db 812 TGTGGATGGCGGCTCCATGTGGGCCCCCGGGCCGTTcAGGTACCAGACCAGATACGAGC 871  
QY 187 ProLeuGlyGluProProAlaThrProAlaAlaValGlyProThrThrMetAlaSerGly 206  
Db 872 CCCCTA-----CCAGCAGCAGCAGCCGCGCAGCCGCGC---TCCTTCGGGGCCCCCGGC 919  
QY 207 GlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGlyAsnAlaSerGly 226  
Db 920 ACAGGGCCACCCCGCA-----CATGCAGCAGAT-----GGGCAGCTATCTGGC 964  
QY 227 AsnTrpHis-----CysAspSerThrTrpLeuGlyAspArgValIleThrThrSer 243  
Db 965 ACGTGGGG-ATTTCTCCATGCAGCAGCAG-----GCGAGCCAC 1002  
QY 244 ThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSerAla 263  
Db 1003 AGCAGAGGATGGC-----CAGTTTCCCAAGGC 1031  
QY 264 SerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyrPhe 283  
Db 1032 CAAGAGGGCCTCAGTCAGGAAGCCCTTTcATTGCCACCTCAGGACCC---GGCCACCTG 1088  
QY 284 AspPhe-----Asn 286  
Db 1089 TCGCACATGCCCCAGCAGAGCCCGCAGATGGCCCTTCCCTGGCCACCCTCAGTCGACGAG 1148  
QY 287 ArgPheHisCysHis-----Phe 292  
Db 1149 CAGTTCCACCATCACCCCGCTGTCTCCACGGAGAAATCCGTTGCCACAGTCCCGATTC 1208  
QY 293 SerProArgAspTrpGlnArgLeuIleAsnAsnAsnTyrGlyPheArgProLysArgLeu 312  
Db 1209 TCCCCCAACCCTCCTCAACA-----GGGGTGTcAGGCCACAAACCCCTT 1253  
QY 313 AsnPheLysLeuPheAsnIleGlnValLysGluValThrThrAsnAspGlyValThrThr 332  
Db 1254 AACTTTAGTTCTCGGAACCAGACAGTCCCTCCCTCTCTGTAAACAAC----- 1301  
QY 333 IleAlaAsnAsnLeuThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuPro 352  
Db 1302 -----TCAGGGCAGTATTCTCGATATCCT 1325  
QY 353 TyrValLeuGlySerAlaHisGlnGlyCysLeu----- 363  
Db 1326 TAC-----AGTAACCTAAATCAGGGATTAGTTAAACAGTACAGGGATGAATCAAAATTTA 1379  
QY 364 -----ProProPheProAlaAspValPheMetIleProGlnTyrGlyTyr 378  
Db 1380 GGCCTTACAAACAGTACTCCAATGAATCAGTCGGTA-----CCAAGATAC----- 1424  
QY 379 LeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyr 398  
Db 1425 -----CCCAATGCCGTGGGA----- 1439  
QY 399 PheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGlu 418  
Db 1440 TTCCCGTCA-----AACAGTGTGTCAGGGACTCGTG-----CACCAG 1475  
QY 419 ValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeu 438  
Db 1476 CAGCCTATCCACTCCAGTGGCTCACTGAACCAAAATGAACACAAACTATGCACCCT--- 1532

QY 439 IleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsn 458  
Db 1533 -----TCACAGCCTCAGGGAACGTACGCCTCTCCACCTCCC 1568  
QY 459 LysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrp 478  
Db 1569 ATGTACCCCATGAAGCAATGAGTAACCCAGCAGGC-----ACGCCGCTCCGCAGGTC 1622  
QY 479 LeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsn 498  
Db 1623 AGGCCCGGAAG-TGCTGG-----GAT 1642  
QY 499 SerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIle 518  
Db 1643 GCCCATGGAAGTTGGCAG--TTATCCAAATAT-----GCCCCACCTCAGCCATC 1690  
QY 519 AsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSer 538  
Db 1691 TCACCAGCCCCCTGGTGCCATGGGAATCGGACAGAGAA----- 1729  
QY 539 GlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnVal 558  
Db 1730 -----TATGGGCCCCAGAAACATGCAGCAGCCTCGTTCTTTATGGGCATGTC 1777  
QY 559 MetIleThrAsp-----GluGluGluIleLys----- 567  
Db 1778 CTCGGCACCCAGGAGCTGACCGGACACATAGAGACCAATGGTTGCCCTGGTGTGGCCT 1837  
QY 568 -----AlaThrAsnProValAlaThrGluArgPheGlyThrValAlaValAsnPhe 584  
Db 1838 TCGGGATCCACAAGCAATCCAGGAACGACTGATACCTGGCCAACAGCACCC----- 1888  
QY 585 GlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeuProGly 604  
Db 1889 TGGTCAGCAGCCATCTTCCAGCAGTTGCCAACCTGTCTCCACTACAGCCCCACCC--- 1945  
QY 605 MetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTyrAlaLysIle-ProHi 624  
Db 1946 -----AGGCTGCACCAGTCTTCACCTCCACATCCCCA 1978  
QY 624 sThrAsp-----GlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAs 642  
Db 1979 TCACCAGCCTTGGGCACAACTCCACCCGTCACCC-----CAGAA 2017  
QY 642 nProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPh 662  
Db 2018 TACCCACAGAAAGTG-----CCTGTGCATCAGCATTCCTCCATCGGAGCC 2062  
QY 662 eSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerVal-- 681  
Db 2063 CTTTCTAGAGAAGCCAGTCCGGATATGACCCAGGTcAGCGCACAGAACGCTCAGCTAGT 2122  
QY 682 -----GluIleGluTrpGluLeuGlnLysGluAsnSerLysAr 694  
Db 2123 GAAGAGCGACGATTACCTGCCGTGCGATAGAGCAGCAGCCACAGCAGAAGAGAGAAAAA 2182  
QY 694 gTrpAsnProGluVal 699  
Db 2183 GAAAAACAACCCACATT 2198

RESULT 7

AK087510 3943 bp mRNA linear HTC 05-DEC-2002  
LOCUS Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched  
DEFINITION library, clone:EI30308C19 product:weakly similar to TESTIS-SPECIFIC  
Y-ENCODED-LIKE PROTEIN (FRAGMENT) [Homo sapiens], full insert  
sequence.  
ACCESSION AK087510  
VERSION AK087510.1 GI:26104328  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



Qy 175 GlnThrGlyAspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThr 194  
Db 515 TCTGCGGGTGTGCGCGTCT-----CCAGTGGGGAGGAAAAAAGATG 562  
Qy 195 ProAlaAlaValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsn 214  
Db 563 -----ACAGAGAAGCATGCTGGGCGAGGCTCTCCCGCGACA----- 598  
Qy 215 AsnGluGlyAlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrp 234  
Db 599 -----GTGGGCAGCATGGAT----- 613  
Qy 235 LeuGlyAspArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsn 254  
Db 614 -----ACCTAGAGACGGTCCAGCTAAAGCTAGAGACCATGAAT--- 652  
Qy 255 HisLeuTyrLysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPhe 274  
Db 653 -----GCACAGGCTGACAGGGCATATCTC 676  
Qy 275 GlyTyrSerThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerPro 294  
Db 677 AGCTTTTCACGCAAGTTTGGCCAGTTG-----CGACTTCAC---CACTTAGAGCGC 724  
Qy 295 ArgAsp-----TrpGlnArgLeuIleAsnAsnAsnTrp 305  
Db 725 CGGAACCTCTCATCCAGAGCATTCGGGGCTTCTGGGGCAAGCTTTTCAGAACCAT--- 781  
Qy 306 GlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLysGluValThr 325  
Db 782 -----CCCCAGCTGTACAGCTTT-----CTGAATACCAAGATAAGGAG----- 820  
Qy 326 ThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGlnValPheSer 345  
Db 821 -----GTATTGAGCTATTGAACAGACTGGAGGTGAAGAACTTGGCCTTGCC 868  
Qy 346 AspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCysLeuProPro 365  
Db 869 AGGTGGGCTACAAAATCAAGTTCTACTTTGGCCGA-----AACCCC 910  
Qy 366 PheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsnAsnGlySer 385  
Db 911 TATTTCAAAACAAGGTGCTCATTAAGGAATATGGC-----TGTGGTCCATCCGGT 961  
Qy 386 GlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGlnMetLeuArg 405  
Db 962 CAAGTGTGTCTCGTTTCAGCT-----CCAATCCAGTGGTCCCGAGCCATGATCTACAA 1015  
Qy 406 Thr-----GlyAsnAsnPheThrPheSerTyrThrPheGluGluVal 419  
Db 1016 TCCCTAAGCAAGGAAACCAGAAAACAATGGTAGCTTCTTTGGGTGTTTTCAAAT--- 1072  
Qy 420 ProPheHisSerSerTyrAlaHisSerGlnSerLeu-----AspArgLeu 434  
Db 1073 -----CACAGCTCTATTGAGTCTGACAAAGATTGTTGAGATAAATAATGAGGACCTGTGG 1126  
Qy 435 MetAsnProLeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGly 454  
Db 1127 CCCAATCTCTA-----CAGTACTACCTGATCATGTAAGAACCCCGTGGAGAGAAAGGA 1180  
Qy 455 SerAlaGlnAsnLysAspLeuPheSerArgGlySerProAlaGlyMetSerValGln 474  
Db 1181 AAGGAAGAA-----AGGCCAGGTCCAGCAAAACTGAGC----- 1213  
Qy 475 ProLysAsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThr 494  
Db 1214 -----CCTGCACCTGCAGTGAGGCGAGCCCACTGATCCGACACAGATCTC 1258  
Qy 495 AspAsnAsnAsnSerAsnPheThrTrpThr 504  
Db 1259 CCTACCACCTCT-----TGCTGGACC 1279

RESULT 8  
CA356072/c  
LOCUS  
DEFINITION

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

FEATURES  
source

1. .640  
/organism="Oncorhynchus mykiss"  
/mol\_type="mRNA"  
/db\_xref="taxon:8022"  
/clone="1RT91G17\_A\_D09"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/clone\_lib="NCCCW 1RT"  
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;  
Library made from pooled tissue from brain, gill, liver,  
spleen, muscle, and kidney."  
BASE COUNT 136 a 185 c 198 g 121 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.818 Length: 640  
Score: 125.00 Matches: 63  
Percent Similarity: 37.05% Conservative: 30  
Best Local Similarity: 25.10% Mismatches: 71  
Query Match: 3.13% Indels: 88  
DB: 14 Gaps: 13

US-09-807-802A-13 (1-736) x CA356072 (1-640)

Qy 124 ArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaProGlyLysLys 143  
Db 602 AGACCTGTTTCCCGCTTGGTCCCATTTGCTCCGGGCAGACCGTGAGGTCCAGGT----- 549  
Qy 144 ArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly---LysThr 162  
Db 548 ATGCCGGTGAAGAGATTCCAGCAGGTCCGGGGCTTCCAGCAGGTCCAGGCATGCCCCCTG 489  
Qy 163 GlyGln---GlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGluSer 181  
Db 488 GGACCCCTGAAGTCCAGCTGCCCCACCTTTCTTTC-----TCGCCAGCT 444  
Qy 182 ValProAsp-----ProGlnProLeuGlyGluPro----- 191  
Db 443 GCACCAGGAACACCAGGCTTACCAGGACCACCACCGGGGAGCCAGGTTTCCAGGTGCGGAG 384  
Qy 192 -----ProAlaThrProAlaAlaValGly---ProThrThrMet 203  
Db 383 TAGCCAGCAGGTCCGGGTGGTCCGGCAGGGCCCTCTGGTCCGGGTGTCCACAGCTCTT 324  
Qy 204 AlaSerGlyGly---GlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 222

CA356072 640 bp mRNA linear EST 05-NOV-2002  
628112 NCCCW 1RT Oncorhynchus mykiss cDNA clone 1RT91G17\_A\_D09 5',  
mRNA sequence.

CA356072  
CA356072.1 GI:24601259  
EST.

Oncorhynchus mykiss (rainbow trout)

Oncorhynchus mykiss  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

1 (bases 1 to 640)

Rexroad, C.E. and Keele, J.W.

Sequence analysis of a rainbow trout normalized cDNA library

Unpublished

Contact: Rexroad CE

USDA, ARS, National Center for Cool and Cold Water Aquaculture

11876 Leetown Road, Kearneysville, WV 25430, USA

Tel: 304 724 8340 x2129

Fax: 304 725 0351

Email: crexroad@nccwa.ars.usda.gov

Single pass sequencing. Bases called with phred v0.020425.c and  
trimmed with the aid of the trim\_alt option. Vector identified by  
cross\_match v0.990329.

Seq primer: ACGGATAACAATTTTCACACAGGA.

Location/Qualifiers



```
Db      323 TCCCTGGGGGCCAGGTGGGCCAGGGGGCCCTCTGGGCCGGGCTCACCTGGGGTA--- 267
QY      223 AenAlaSerGlyAsnTrpHisCysAspSerThrTrpLeu----- 235
Db      266 -----CCGGGCTCTCCTGCCACTGACACCACGTCGCTCTTCACAGAGTAGGGCTGGTAC 213
QY      236 ---GlyAspArgValIleThr-ThrSerThrArgThrTrpAlaLeuProThrTyrAsnAs 254
Db      212 TGGGGGAGGCCCTTCATTACTTCTTCA-----CTATATATGAA 174
QY      254 nHisLeuTyrLysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPh 274
Db      173 CCACTACCATGAACAGCTGCA----- 152
QY      274 eGlyTyrSerThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerPr 294
Db      151 -----AGCCC 147
QY      294 oArgAspTrpGlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPh 314
Db      146 ACCAGGAGG-AGGAGGATGCTCAATACT-----CTTAC 115
QY      314 eLysLeuPheAsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAl 334
Db      114 TTCCATTTCAGCATGCAGTTGAGTGAGTTGGAGTTCCAGAAGGGGTAAGAAATCTTCA 55
QY      334 aAsnAsnLeuThrSerThrValGlnValPhe 344
Db      54 GCACCACTGTCTCTCTACTGCTGCTGAGTTTAT 24
```

RESULT 9  
CA363789/c  
LOCUS  
DEFINITION  
638563 NCCCWA 1RT Oncorhynchus mykiss cDNA clone 1RT118G06\_C\_D03  
5', mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
CA363789.1 GI:24672376  
Oncorhynchus mykiss (rainbow trout)  
Oncorhynchus mykiss  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
1 (bases 1 to 701)  
Rexroad, C.E. and Keele, J.W.  
Sequence analysis of a rainbow trout normalized cDNA library  
Unpublished  
Contact: Rexroad CE

USDA, ARS, National Center for Cool and Cold Water Aquaculture  
11876 Leetown Road, Kearneysville, WV 25430, USA  
Tel: 304 724 8340 x2129  
Fax: 304 725 0351

Email: crexroad@nccwa.ars.usda.gov  
Single pass sequencing. Bases called with phred v0.020425.c and  
trimmed with the aid of the trim\_alt option. Vector identified by  
cross match v0.990329.

Seq primer: AGCGGATAACAATTTTCACACAGGA.

Location/Qualifiers

FEATURES  
source  
1..701  
/organism="Oncorhynchus mykiss"  
/mol\_type="mRNA"  
/db\_xref="taxon:8022"  
/clone="1RT118G06\_C\_D03"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/clone\_lib="NCCCWA 1RT"  
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;  
Library made from pooled tissue from brain, gill, liver,  
spleen, muscle, and kidney."

BASE COUNT  
ORIGIN  
150 a 194 c 231 g 126 t

Alignment Scores:  
Pred. No.: 0.927 Length: 701  
Score: 125.00 Matches: 63  
Percent Similarity: 37.05% Conservative: 30  
Best Local Similarity: 25.10% Mismatches: 71  
Query Match: 3.13% Indels: 88  
DB: 14 Gaps: 13

US-09-807-802A-13 (1-736) x CA363789 (1-701)

```
QY      124 ArgValLeuGluProLeuGlyLeuValGluGlyAlaLysThrAlaProGlyLysLys 143
Db      585 AGACCTGTTTCCCCCTTGGTCCCATTTGCTCCGGGCAGACCGTGAGGTCCAGGT----- 532
QY      144 ArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly---LysThr 162
Db      531 ATGCCGTTGGAAGAGATTCCAGCAGGTCCGGGGCTTCCAGCAGGTCCAGGCATGCCCTGT 472
QY      163 GlyGln---GlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGluSer 181
Db      471 GGACCCCTGAAGTCCAGTCCCCCACCCTTCTCTTC-----TCGCCAGCT 427
QY      182 ValProAsp-----ProGlnProLeuGlyGluPro----- 191
Db      426 GCACCAGGAACACCAGGCTTACCAGGACCAACCCGGGAGCCAGGTTTGCAGGTGCGGAG 367
QY      192 -----ProAlaThrProAlaAlaValGly---ProThrThrMet 203
Db      366 TAGCCAGCAGTCCGGGTGGTCCGGCAGGCGCCCTCTGGTCCGGGCTGTCCACAGCTCTT 307
QY      204 AlaSerGlyGly---GlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 222
Db      306 TCCCTGGGGGCCAGGTGGCCAGGGGGCCCTCTGGGCGGGCTCACTGGGGTA--- 250
QY      223 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeu----- 235
Db      249 -----CCGGGCTCTCTCTGCCACTGACACACCGTGGCTCTTTCACAGAGTAGGGCTGGTAC 196
QY      236 ---GlyAspArgValIleThr-ThrSerThrArgThrTrpAlaLeuProThrTyrAsnAs 254
Db      195 TGGGGGAGGCCCTTCATTACTTCTTCA-----CTATATATGAA 157
QY      254 nHisLeuTyrLysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPh 274
Db      156 CCACTACCATGAACAGCTGTCA----- 135
QY      274 eGlyTyrSerThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerPr 294
Db      134 -----AGCCC 130
QY      294 oArgAspTrpGlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPh 314
Db      129 ACCAGGAGG-AGGAGGATGCTCAATACT-----CTTAC 98
QY      314 eLysLeuPheAsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAl 334
Db      97 TTCCATTTCAGCATGCAGTTGAGTGAGTTGGAGTTCCAGAAGGGGTAAGAAATCTTCA 38
QY      334 aAsnAsnLeuThrSerThrValGlnValPhe 344
Db      37 GCACCACTGTCTCTCTACTGCTGCTGAGTTTAT 7
```

RESULT 10

CA378088/c

LOCUS

DEFINITION

CA378088

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

CA378088 703 bp mRNA linear EST 06-NOV-2002  
656813 NCCCWA 1RT Oncorhynchus mykiss cDNA clone 1RT42114\_C\_E07 5',  
mRNA sequence.

CA378088

CA378088.1 GI:24697641

EST.

Oncorhynchus mykiss (rainbow trout)

Oncorhynchus mykiss

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus. 1 (bases 1 to 703)

Rexroad,C.E. and Keele,J.W.

Sequence analysis of a rainbow trout normalized cDNA library

Unpublished

Contact: Rexroad CE

USDA, ARS, National Center for Cool and Cold Water Aquaculture 11876 Leetown Road, Kearneysville, WV 25430, USA

Tel: 304 724 8340 x2129

Fax: 304 725 0351

Email: crexroad@cccwa.ars.usda.gov

Single pass sequencing. Bases called with phred v0.020425.c and trimmed with the aid of the trim\_alt option. Vector identified by cross\_match v0.990329.

Seq primer: AGCGGATAACAAATTTACACAGGA.

Location/Qualifiers

1..703

/organism="Oncorhynchus mykiss"

/mol\_type="mRNA"

/db\_xref="taxon:8022"

/clone="1RT42I14\_C\_E07"

/tissue\_type="pooled"

/lab\_host="DH10B"

/clone\_lib="NCCCWALRT"

/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI; Library made from pooled tissue from brain, gill, liver, spleen, muscle, and kidney."

BASE COUNT 146 a 195 c 235 g 127 t

ORIGIN

Alignment Scores:

Pred. No.: 0.93 Length: 703

Score: 125.00 Matches: 54

Percent Similarity: 41.04% Conservative: 17

Best Local Similarity: 31.21% Mismatches: 55

Query Match: 3.13% Indels: 47

DB: 14 Gaps: 11

US-09-807-802A-13 (1-736) x CA378088 (1-703)

Qy 124 ArgValLeuGluProLeuGlyLeuValGluGluGlyAlaLysThrAlaProGlyLysLys 143

Db 578 AGACCTGTTTCCCCCTTGGTCCCATTTGCTCCGGGCAGACCGTGAGGTCCAGGT----- 525

Qy 144 ArgProValGluGlnSerProGlnGluProAspSerSerSerGlyIleGly---LysThr 162

Db 524 ATGCCGGTGAAGAGATTCCAGCAGGTCCGGGGCTTCCAGCAGGTCCAGGCATGCCCTG 465

Qy 163 GlyGln---GlnProAlaLysLysArgLeuAsnPheGlyGlnThrGlyAspSerGluSer 181

Db 464 GGACCCCTGAAGTCCAGCTGCCCCACCTTTCCTTTC-----TCGCCAGCT 420

Qy 182 ValProAsp-----ProGlnProLeuGlyGluPro----- 191

Db 419 GCACCAGGAACACACAGGCTTACCAGGACCACCGGGGAGCCAGGTTTGCCAGGTGCGGAG 360

Qy 192 -----ProAlaThrProAlaAlaValGly---ProThrThrMet 203

Db 359 TAGCCAGCAGGTCCGGGTGGTCCGGCAGGSCCTCTGCTCCGGGCTGTCCACAGCTCTT 300

Qy 204 AlaSerGlyGly---GlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 222

Db 299 TCCCTGGGGGCCAGGTGGGCCAGGGGGGCCCTCTGGGGCCGGGCTCACCTGGGTA--- 243

Qy 223 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeu----- 235

Db 242 -----CCGGGTCTCTCTGCCACTGACACACAGTGGCTCTTCACAGAGTAGGGCTGTAC 189

Qy 236 ---GlyAspArgValIleThr-ThrSerThrArgThrTrpAlaLeuProThrTyrAsnAs 254

Db 188 TGGGGGAGGCGCTTCATTACCTTCTTCA-----CTATATATGAA 150

Qy 254 nHisLeuTyrLysGlnIleSerSerAlaSerThrGly 266

Db 149 CCACTACCATGAACAGCTGTCAAGCCCAAGGAGGA 113

RESULT 11

BJ139005/c

LOCUS

DEFINITION

Caenorhabditis elegans cDNA clone yk1141g12 3', mRNA sequence.

ACCESSION

BJ139005

VERSION

BJ139005.1 GI:18299171

KEYWORDS

EST.

SOURCE

ORGANISM

Caenorhabditis elegans

Caenorhabditis elegans

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea

; Rhabditidae; Peloderinae; Caenorhabditis.

1 (bases 1 to 745)

REFERENCE

AUTHORS

Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y. and Sugano,S.

A complementary view of the C.elegans genome

TITLE

JOURNAL

COMMENT

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1..745

/organism="Caenorhabditis elegans"

/mol\_type="mRNA"

/strain="N2"

/db\_xref="taxon:6239"

/clone="yk1141g12"

/sex="hermaphrodite"

/tissue\_type="whole animal"

/dev\_stage="L1"

/clone\_lib="unpublished oligo-capped cDNA library, C. elegans L1 stage"

BASE COUNT 75 a 212 c 238 g 216 t 4 others

ORIGIN

Alignment Scores:

Pred. No.: 1.01 Length: 745

Score: 125.00 Matches: 56

Percent Similarity: 35.11% Conservative: 10

Best Local Similarity: 29.79% Mismatches: 72

Query Match: 3.13% Indels: 50

DB: 12 Gaps: 9

US-09-807-802A-13 (1-736) x BJ139005 (1-745)

Qy 87 AspAsnProTyrLeuArgTyrAsnHisAla---AspAlaGluPheGlnGluArgLeu--- 104

Db 678 GAAACACCGGATCCCAAGGAGGATCATGCTNTGGATGCTGCTTCCAGGAGCCGCTGGAC 619

Qy 105 -----GlnGluAspThrSerPheGly 111

Db 618 CAGCCGGAACCCAGGAAGCCAGGAAGACCAGGACGCTCCAGGAGCAGCTGGACTTCCAA 559

Qy 112 GlyAsnLeuGlyArgAlaValPheGlnAlaLysLysArgValLeu---GluProLeuGly 130

Db 558 GGAACCCAGGACGCCCCACAGCACAGCCATGTGAGCCAATCACCCACCACCATGCAAG 499

Qy 131 LeuValGluGluGlyAlaLysThrAlaProGlyLysLysArgProValGluGln----- 148

Db 498 CCATGCCCCAAGGACCCAGCTGGAGCCCCCAGGAGCCCCAGGACCAAGGAGATGCCGGA 439

Qy 149 SerProGlnGluProAspSerSerSerGlyIleGlyLysThrGlyGlnGlnProAlaLys 168

Db 438 GCACCAGGACAAGCCGGAACAAGGATCAGGAGCCGAGCCGCCAGGA-----CCAGCTGGA 385

QY 169 LysArg-----LeuAsnPheGlyGlnThrGly----- 177  
:::  
Db 384 CCAAAGGAGCCCGAGAGCCCGAGGAACCCAGGACAAGCCGAGCCCGAGGACCA 325  
QY 178 -----AspSerGluSerValPro----- 183  
Db 324 GGATCCGACGCTCAATCCGAGTCTTCTCCAGGAGCCCGAGGACCAAGCCGAGCACAAGGA 265  
QY 184 AspProGlnProLeuGlyGluPro-----ProAlaThrProAlaAlaVal 198  
Db 264 CCACGAGGACCACTGGATCCCGAGGAGCAGGAGGACCAAGCAAGCCGAGCCCGCA 205  
QY 199 GlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGlyAla 218  
Db 204 GGACCAAAAGGGACCATCAGGAGCCCGAGGAGCAGGAGCCGACGGAACCCAGGAGCC 145  
QY 219 AspGlyValGlyAsnAlaSerGly 226  
Db 144 CCAGGACAGCCAGGACAATCTGGA 121

RESULT 12  
BG207226 754 bp mRNA linear EST 21-APR-2001  
LOCUS RST26694 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.  
DEFINITION BG207226  
ACCESSION BG207226  
VERSION BG207226.1 GI:13728913  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 754)  
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,  
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.,  
Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith  
,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher  
,J., Danzig,J. and Ducar,M.

TITLE Creation of genome-wide protein expression libraries using random  
activation of gene expression  
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)  
MEDLINE 21227151  
PUBMED 11329013

COMMENT Contact: Scott J. Cain  
Athersys, Inc.  
3201 Carnegie Ave, Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: scain@atersys.com  
High quality sequence stop: 326.  
Location/Qualifiers  
1..754  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_line="HT1080"  
/clone\_lib="Athersys RAGE Library"  
/note="See 'Creation of Genome-wide Protein Expression  
Libraries using Random Activation of Gene Expression',  
Nature Biotechnology, in press. Note that even though the  
cell type indicated is HT1080, since a random activation  
method was used, these sequence tags are not necessarily  
expressed in HT1080 under normal circumstances."

FEATURES  
source  
1..754  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_line="HT1080"  
/clone\_lib="Athersys RAGE Library"  
/note="See 'Creation of Genome-wide Protein Expression  
Libraries using Random Activation of Gene Expression',  
Nature Biotechnology, in press. Note that even though the  
cell type indicated is HT1080, since a random activation  
method was used, these sequence tags are not necessarily  
expressed in HT1080 under normal circumstances."

BASE COUNT 223 a 224 c 191 g 115 t 1 others  
ORIGIN

Alignment Scores:  
Pred. No.: 1.02 Length: 754  
Score: 125.00 Matches: 50  
Percent Similarity: 33.06% Conservative: 31  
Best Local Similarity: 20.41% Mismatches: 68  
Query Match: 3.13% Indels: 96  
DB: 10 Gaps: 11

US-09-807-802A-13 (1-736) x BG207226 (1-754)  
QY 444 TyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeuPhe 463  
Db 206 TTTTATCTCAACCAACCGAA-----GCAGGAGTTAAATCTAAGGAGCACTGTC 256  
QY 464 SerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyProCys 483  
Db 257 CAAAGAAAGTACAGAAACAGGCCTGGCCGTGGAATGAGGAACCTGGATG----- 304  
QY 484 TyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThrTrp 503  
Db 305 -----ACTCGACAGGCAAGCCGAGAGTCTACAGATGGTAGCATGAACAGCTACAGTCA 358  
QY 504 ThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThrAla 523  
Db 359 CAAGGA-----AATCTGATTTTCCCTGGGGTTTCAC 388  
QY 524 MetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIlePhe 543  
Db 389 TTGGCCTCTGATAGCCAGCCAGCGATTCCTCGATGGCTTGGCCTGCTCAGCTAGTG 448  
QY 544 GlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAspGlu 563  
Db 449 GGACGCCAGACTCTG----- 463  
QY 564 GluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaValAsn 583  
Db 463 ----- 463.  
QY 584 PheGlnSerSerSerThrAspProAlaThrGlyAspValHis---AlaMetGlyAlaLeu 602  
Db 464 -----ACAACACTGCAATGGGGACATTCACGTAGGAATGATGGCAAC 508  
QY 603 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 622  
Db 509 AAGGGACAGCTG-----GACGCATACTCATCCGGGCCC-----GCGCCACTG 550  
QY 623 Pro-----HisThrAspGlyHisPheHis----- 630  
Db 551 CCGAAAAAACCCAGCTACAAGACACTGCCACCACCGCTGTCAAAGTAGGCCTACTACACAA 610  
QY 631 -----ProSerProLeuMetGlyGlyPhe 638  
Db 611 CGGAAGCCGGCCCGAGGCAAAAAAACCCCGATCGCCGCGACCCCTAGCGCCA----- 664  
QY 639 GlyLeuLysAsnProProProGlnIleLeuIleLysAsnThr-----ProVal 654  
Db 665 TAAATATCCCGACCCACCCACGATCAAGACAAACCCGACCGACCGAGCAACCCACG 724  
QY 655 ProAlaAsnProPro 659  
Db 725 CCGTGGCAACCTCCG 739

RESULT 13  
AK014322  
LOCUS  
DEFINITION

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE

AK014322 2019 bp mRNA linear HTC 05-DEC-2002  
Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length  
enriched library, clone:322402004 product:similar to LAF-4 PROTEIN  
(LYMPHOID NUCLEAR PROTEIN RELATED TO AP4) [Homo sapiens], full  
insert sequence.

AK014322  
AK014322.1 GI:12852094  
HTC; CAP trapper.  
Mus musculus (house mouse)

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1  
Carninci,P. and Hayashizaki,Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253











GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 14, 2004, 15:19:20 ; Search time 6088.71 Seconds  
(without alignments)  
4024.639 Million cell updates/sec

Title: US-09-807-802A-15  
Perfect score: 3251  
Sequence: 1 TAPGKRPVEQSPQEPDSSS.....NNGLYTEPRPIGTRYLTRPL 599

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09807802/runat\_11022004\_175607\_15929/app\_query.fasta\_1.2389  
-DB=GenEmbl -QFMT=fastap -SUFFIX=xge -MINMATCH=0.1 -LOOFC=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09807802@cgn 1 1 12216@runat\_11022004\_175607\_15929 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:\*

1: gb\_ba:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vi:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

| Result No. | Score  | Query Match % | Length | DB | ID       | Description        |
|------------|--------|---------------|--------|----|----------|--------------------|
| 1          | 3251   | 100.0         | 4718   | 14 | AF063497 | AF063497 Adeno-ass |
| 2          | 3229   | 99.3          | 4683   | 14 | AF028704 | AF028704 Adeno-ass |
| 3          | 2832   | 87.1          | 4722   | 14 | AF028705 | AF028705 Adeno-ass |
| 4          | 2815   | 86.6          | 4726   | 14 | AVU48704 | U48704 Adeno-assoc |
| 5          | 2801   | 86.2          | 8179   | 6  | AR264580 | AR264580 Sequence  |
| 6          | 2764.5 | 85.0          | 4072   | 6  | AR222046 | AR222046 Sequence  |
| 7          | 2764.5 | 85.0          | 4072   | 6  | AX205074 | AX205074 Sequence  |
| 8          | 2764.5 | 85.0          | 4679   | 6  | AX282480 | AX282480 Sequence  |
| 9          | 2764.5 | 85.0          | 4679   | 14 | AF043303 | AF043303 Adeno-ass |
| 10         | 2764.5 | 85.0          | 7327   | 12 | AF369963 | AF369963 Cloning v |
| 11         | 2764.5 | 85.0          | 7557   | 6  | AR222045 | AR222045 Sequence  |
| 12         | 2764.5 | 85.0          | 7557   | 6  | AX205073 | AX205073 Sequence  |
| 13         | 2764.5 | 85.0          | 8698   | 6  | AR222044 | AR222044 Sequence  |
| 14         | 2764.5 | 85.0          | 8698   | 6  | AX205072 | AX205072 Sequence  |
| 15         | 2761.5 | 84.9          | 2208   | 14 | AY243021 | AY243021 Non-human |
| 16         | 2735   | 84.1          | 2211   | 14 | AY242997 | AY242997 Non-human |
| 17         | 2722.5 | 83.7          | 4680   | 6  | AR028767 | AR028767 Sequence  |
| 18         | 2722.5 | 83.7          | 4680   | 6  | I62303   | I62303 Sequence 1  |
| 19         | 2715   | 83.5          | 2217   | 14 | AY243009 | AY243009 Non-human |
| 20         | 2710   | 83.4          | 2217   | 14 | AY243015 | AY243015 Non-human |
| 21         | 2709.5 | 83.3          | 4721   | 14 | AF513851 | AF513851 Adeno-ass |
| 22         | 2706.5 | 83.3          | 4675   | 6  | AX135805 | AX135805 Sequence  |
| 23         | 2706.5 | 83.3          | 4675   | 6  | AX286292 | AX286292 Sequence  |
| 24         | 2706.5 | 83.3          | 4675   | 6  | BD094552 | BD094552 Method of |
| 25         | 2706.5 | 83.3          | 4675   | 14 | AA2CG    | J01901 Adeno-assoc |
| 26         | 2703   | 83.1          | 2187   | 14 | AY243013 | AY243013 Non-human |
| 27         | 2701.5 | 83.1          | 2196   | 14 | AY243010 | AY243010 Non-human |
| 28         | 2700   | 83.1          | 2217   | 14 | AY243007 | AY243007 Non-human |
| 29         | 2699   | 83.0          | 2217   | 14 | AY243023 | AY243023 Non-human |
| 30         | 2696.5 | 82.9          | 2208   | 14 | AY243008 | AY243008 Non-human |
| 31         | 2696.5 | 82.9          | 2214   | 14 | AY243020 | AY243020 Non-human |
| 32         | 2696   | 82.9          | 2187   | 14 | AY243006 | AY243006 Non-human |
| 33         | 2695   | 82.9          | 2187   | 14 | AY243017 | AY243017 Non-human |
| 34         | 2692   | 82.8          | 2217   | 14 | AY243022 | AY243022 Non-human |
| 35         | 2691.5 | 82.8          | 2196   | 14 | AY243011 | AY243011 Non-human |
| 36         | 2689   | 82.7          | 2187   | 14 | AY243016 | AY243016 Non-human |
| 37         | 2686   | 82.6          | 2187   | 14 | AY243019 | AY243019 Non-human |
| 38         | 2684.5 | 82.6          | 2190   | 14 | AY242998 | AY242998 Non-human |
| 39         | 2684.5 | 82.6          | 2202   | 14 | AY243014 | AY243014 Non-human |
| 40         | 2683.5 | 82.5          | 2202   | 14 | AY243012 | AY243012 Non-human |
| 41         | 2682.5 | 82.5          | 2202   | 14 | AY243004 | AY243004 Non-human |
| 42         | 2679.5 | 82.4          | 2190   | 14 | AY242999 | AY242999 Non-human |
| 43         | 2679   | 82.4          | 4393   | 14 | AF513852 | AF513852 Adeno-ass |
| 44         | 2676.5 | 82.3          | 2190   | 14 | AY243000 | AY243000 Non-human |
| 45         | 2674   | 82.3          | 2187   | 14 | AY243018 | AY243018 Non-human |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ALIGNMENTS

AF063497  
LOCUS AF063497 4718 bp DNA linear VRL 27-APR-1999  
DEFINITION Adeno-associated virus 1, complete genome.  
ACCESSION AF063497  
VERSION AF063497.1 GI:4689096  
KEYWORDS  
SOURCE Adeno-associated virus 1  
ORGANISM Adeno-associated virus 1  
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.  
REFERENCE 1 (bases 1 to 4718)  
AUTHORS Xiao,W., Chirmule,N., Berta,S.C., McCullough,B., Gao,G. and Wilson,J.M.  
TITLE Gene therapy vectors based on adeno-associated virus type 1  
JOURNAL J. Virol. 73 (5), 3994-4003 (1999)  
MEDLINE 99214338  
PUBMED 10196295  
REFERENCE 2 (bases 1 to 4718)  
AUTHORS Xiao,W. and Wilson,J.M.  
TITLE Direct Submission  
JOURNAL Submitted (05-MAY-1998) IHGT, University of Pennsylvania, 3601 Spruce Street, Philadelphia, PA 19104, USA  
FEATURES  
source  
1..4718  
/organism="Adeno-associated virus 1"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:85106"  
/note="AAV1"  
335..2206  
/codon\_start=1  
/product="nonstructural protein"  
/protein\_id="AAD27758.1"  
/db\_xref="GI:4689098"  
/translation="MPGFYEIVIKVPSDLDEHLPGISDSFVSWAEKEWELPPDSMD  
LNLEQAPLTVAEKLQDFLVQRRVSKAPEALFFVQFEKGESYFHLHLIVETGVKS  
MVLGRFLSQIRDLKVQTIYRGIEPTLPNWFAVTKTRNGAGGKVVDECYIPNYLLPK  
TQELQWANTWMEEYISACNLAEKRLVAQHLTHVSQTQEQNKENLNPNSDAPVIRS  
KTSARYMELVGLVDRGITSKQWIQEDQASYISFNAASNSRSQIKAAALDNAGKIMAL  
TKSAPDYLVGPPADIKTNRIYRILELNGYEPAYAGSVFLGWAQKRFGRKNTIWLFG  
PATTKGTNIAEAIHAHVPFGCVNWTNENFPFNDQVDMQVWEEGKMTAKVYESAKA  
ILGSKVRVDQCKSSAQIDPTPIVTSNTMCAVIDGNSTTFEHPQQLQDRMFKFEL  
TRLEHDFGKVTQEVKFFRWAQDHVTEVAHEFYVRKGANKRPAPDDADKSEPKRA  
CPSVADPSTSDAEGAPVDFADRYQNKCSRHAGMLQMLFPCKTCERMNQNFICFTHGT  
RDCSECFPGVSESQPVVRKRTYRKLCIAIHLLGRAPEIACSDLVNVVDLDDCVSEQ"  
2223..4433  
/codon\_start=1  
/product="capsid protein"  
/protein\_id="AAD27757.1"  
/db\_xref="GI:4689097"

CDS

1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20  
2634 ACGGCTCCTGGAAAGAAACGTCGCGTAGAGAGTCGCCACAAGAGCCAGACTCCTCCTCG 2693  
21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40  
2694 GGCATCGGCAAGACAGGCCAGCAGCCCGCTAAAGAGAGACTCAATTTTGGTCAGACTGGC 2753  
41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60  
2754 GACTCAGAGTCAGTCCCGGATCCACAACCTCTCGAGAGAACTCCAGCAACCCCGCTGCT 2813  
61 ValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGly 80  
2814 GTGGGACCTACTACATGGCTTCAGGCGGTGGCGCAACCAATGGCAGACAAATAACGAAGGC 2873  
81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100  
2874 GCCGACGGAGTGGGTAAATGCCTCAGGAAATTTGGCATTCGATTCCACATGGTGGCGAC 2933  
101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTrpAsnAsnHisLeuTyr 120  
2934 AGAGTCATCACCAACCAACACCCGACCTGGGCCCTGGCCACCTACATAAACCACTCTAC 2993  
121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140  
2994 AAGCAATCTCCAGTGTCTCAACGGGGGCCAGCAACGACCAACCACTACTTCGGCTACAGC 3053  
141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160  
3054 ACCCCCTGGGGTATTTTGATTTCAACAGATTCCACTGCCACTTTTCACCACGTGACTGG 3113  
161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180  
3114 CAGCGACTCATCAACAACAATTTGGGATTTCGGGCCCAAGAGACTCAACTTCAAACTCTTC 3173  
181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200  
3174 AACATCCAAGTCAAGGAGGTACGACGAATGATGGCGTCACAACCATCGCTAATAACCTT 3233  
201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220  
3234 ACCAGCAGGTTCAAGTCTTCTCGGACTCGGAGTACAGAGCTTCGTACGTCTCGGCTCT 3293  
221 AlaHisGlnGlyCysLeuProProPheProAlaAlaAppValPheMetIleProGlnTyrGly 240  
3294 GCGCACAGGGCTGCCCTCCCTCCGTTCCCGGGGACGTGTTTCATGATTCGCAATACGGC 3353  
241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260  
3354 TACCTGACGCTCAACAATGGCAGCCCAAGCCGTGGAGCTTCATCTTTTACTGCTCGGAA 3413  
261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280  
3414 TATTTCCCTTCTCAGATGCTGAGAACGGGCAACAACCTTACCTTCAGCTACACCTTTGAG 3473  
281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300  
3474 GAAGTGCCTTTCCACAGCAGTACGCGCACAGCCAGCGCTGGACCGCTGATGAATCCT 3533  
301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320  
3534 CTCATCGACCAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCTCA 3593  
321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340  
3594 AACAAAGGACTTGCTGTTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTACAGCCCAAAAC 3653  
341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360  
3654 TGGCTACCTGGACCTGTATTATCGGCAGCAGCGCGTTTCTAAAAACAAAAACAGACACAAC 3713

CDS

1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20  
2634 ACGGCTCCTGGAAAGAAACGTCGCGTAGAGAGTCGCCACAAGAGCCAGACTCCTCCTCG 2693  
21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40  
2694 GGCATCGGCAAGACAGGCCAGCAGCCCGCTAAAGAGAGACTCAATTTTGGTCAGACTGGC 2753  
41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60  
2754 GACTCAGAGTCAGTCCCGGATCCACAACCTCTCGAGAGAACTCCAGCAACCCCGCTGCT 2813  
61 ValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGly 80  
2814 GTGGGACCTACTACATGGCTTCAGGCGGTGGCGCAACCAATGGCAGACAAATAACGAAGGC 2873  
81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100  
2874 GCCGACGGAGTGGGTAAATGCCTCAGGAAATTTGGCATTCGATTCCACATGGTGGCGAC 2933  
101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTrpAsnAsnHisLeuTyr 120  
2934 AGAGTCATCACCAACCAACACCCGACCTGGGCCCTGGCCACCTACATAAACCACTCTAC 2993  
121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140  
2994 AAGCAATCTCCAGTGTCTCAACGGGGGCCAGCAACGACCAACCACTACTTCGGCTACAGC 3053  
141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160  
3054 ACCCCCTGGGGTATTTTGATTTCAACAGATTCCACTGCCACTTTTCACCACGTGACTGG 3113  
161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180  
3114 CAGCGACTCATCAACAACAATTTGGGATTTCGGGCCCAAGAGACTCAACTTCAAACTCTTC 3173  
181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200  
3174 AACATCCAAGTCAAGGAGGTACGACGAATGATGGCGTCACAACCATCGCTAATAACCTT 3233  
201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220  
3234 ACCAGCAGGTTCAAGTCTTCTCGGACTCGGAGTACAGAGCTTCGTACGTCTCGGCTCT 3293  
221 AlaHisGlnGlyCysLeuProProPheProAlaAlaAppValPheMetIleProGlnTyrGly 240  
3294 GCGCACAGGGCTGCCCTCCCTCCGTTCCCGGGGACGTGTTTCATGATTCGCAATACGGC 3353  
241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260  
3354 TACCTGACGCTCAACAATGGCAGCCCAAGCCGTGGAGCTTCATCTTTTACTGCTCGGAA 3413  
261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280  
3414 TATTTCCCTTCTCAGATGCTGAGAACGGGCAACAACCTTACCTTCAGCTACACCTTTGAG 3473  
281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300  
3474 GAAGTGCCTTTCCACAGCAGTACGCGCACAGCCAGCGCTGGACCGCTGATGAATCCT 3533  
301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320  
3534 CTCATCGACCAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCTCA 3593  
321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340  
3594 AACAAAGGACTTGCTGTTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTACAGCCCAAAAC 3653  
341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360  
3654 TGGCTACCTGGACCTGTATTATCGGCAGCAGCGCGTTTCTAAAAACAAAAACAGACACAAC 3713

BASE COUNT 1121 a 1393 c 1273 g 931 t  
ORIGIN

Alignment Scores:

Pred. No.: 4, 83e-231 Length: 4718  
Score: 3251.00 Matches: 599  
Percent Similarity: 100.00% Conservative: 0  
Best local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-09-807-802A-15 (1-599) x AF063497 (1-4718)

```
Qy 361 AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380
Db 3714 AACAGCAATTTTACCCTGGACTGGTGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATC 3773
Qy 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400
Db 3774 ATCAACCTTGGCACTGCTATGGCCTCACACAAAGACGACGAAGACAAGTTCTTCCCATG 3833
Qy 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420
Db 3834 AGCGGTGTCATGATTTTGGAAAAGAGAGCGCGCGAGCTTCAACACTGCATTGGACAAT 3893
Qy 421 ValMetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440
Db 3894 GTCATGATTACAGACGAAGAGGAATTAAGCCACTAAACCCTGTGGCCACCAGAAAGATT 3953
Qy 441 GlyThrValAlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHis 460
Db 3954 GGGACCGTGGCAGTCAATTTCCAGAGCAGCAGCACGACCCCTCGACCGGAGATGTGCAT 4013
Qy 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 480
Db 4014 GCTATGGGAGCATTACCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGACGGTCCC 4073
Qy 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500
Db 4074 ATTTGGGCCAAAATTCCTCACACAGATGGACATTTTCAACCGTCTCCTCTTATGGCGGC 4133
Qy 501 PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520
Db 4134 TTTGGACTCAAGAACCCGCCTCCTCAGATCCTCATCAAAAACACGCCTGTTCTTCGCAAT 4193
Qy 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540
Db 4194 CTTCCGGCGGAGTTTTCAGCTACAAAGTTTGCTTCAATTCATCACCCAATACCTCCACAGGA 4253
Qy 541 GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro 560
Db 4254 CAAGTGAGTGTGGAATTGAATGGGAGCTGCAGAAAGAAAACAGCAAGCGTGGATCCC 4313
Qy 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580
Db 4314 GAAGTGCAGTACACATCCAAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAC 4373
Qy 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599
Db 4374 AATGGACTTTATACTGAGCCTCGCCCCATTGGCACCCGTTACCTTACCCGTCCCCCTG 4430

RESULT 2
AF028704
LOCUS AF028704 4683 bp DNA linear VRL 12-JAN-1998
DEFINITION Adeno-associated virus 6, complete genome.
ACCESSION AF028704
VERSION AF028704.1 GI:2766605
KEYWORDS
SOURCE Adeno-associated virus 6
ORGANISM Adeno-associated virus 6
REFERENCE 1 (bases 1 to 4683)
AUTHORS Rutledge,E.A., Halbert,C.L. and Russell,D.W.
TITLE Infectious clones and vectors derived from adeno-associated virus
(AAV) serotypes other than AAV type 2
JOURNAL J. Virol. 72 (1), 309-319 (1998)
MEDLINE 98080418
PUBMED 9420229
REFERENCE 2 (bases 1 to 4683)
AUTHORS Rutledge,E.A. and Russell,D.W.
TITLE Direct Submission
JOURNAL Submitted (24-SEP-1997) Hematology, Univ. of Washington, Box
357720, Seattle, WA 98195, USA
FEATURES
location/Qualifiers
1..4683
/organism="Adeno-associated virus 6"
```

```
/mol_type="genomic DNA"
/db_xref="taxon:68558"
/note="AAV6"
320..2191
/codon_start=1
/evidence=not_experimental
/product="nonstructural protein Rep78"
/protein_id="AAB95449.1"
/db_xref="GI:2766606"
/translation="MPGFVEIVIKVPSDLDEHLPGISDSFVNWVAEKBWELPPDSDMD
LNLIEQAPLTVAEKLQRDELVQWRRVSKAPEALFFVQFEKGSYFHLHLVETTVGKVS
MVLGRFLSQIRDLKLVQTIYRGIEPTLPNWFVATKTRNGAGGNGKVVDCEYIPNYLLPK
TQPELQWAWTNMEEYISACNLAEKRKLVAHDLTHVSQTQONKENLNPNSDAPVIRS
KTSARYMELVGLVDRGITSEKQWQIEDQASYISFNAASNSRSQIKAALDNAGKIMAL
TKSAPDYLVGPPAPPADIKTNRIYRILELNGYDPAAYAGSVFLGMAQKRFGRNTIWLFG
PATCKTNIAEAIHAHAVPFYGCNVNTNENFPENDCVDMVIWEEGKMTAKVESA
ILGSKVRVVDQCKSSAQIDPTPVIIVTSNTNMCVIDGNSITFEHQQLQDRMFKFEL
TRRLEHDFGKVTQKEVKEFFRWAQDHVTEVAHEFYVRKGGANKRPAPDDADKSEPKRA
CPSVADPSTSDAEGAPVDFADRYQNKCSRHAGMLQMLFPCKTCERMNQNFNICFTHTG
RDCSECFPGVSESQPVVRKTYRKLCIAIHLLGRAPEIACSACDLVNVLDLDDCVSEQ"
2208..4418
/codon_start=1
/evidence=not_experimental
/product="capsid protein VP1"
/protein_id="AAB95450.1"
/db_xref="GI:2766607"
/translation="MAADGYLPDWLEDNLSEGIREWWDLKPAGPKPKANQQKQDDGRG
LVLPGYKYLGPFGNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEPQE
RLQEDTSFGGNLGRAVFAQKRVLEFPFGLVEEGAKTAPGKRPVQSPQEPDSSSGIG
KTQOPAKKRLNFGQTDSESVDPQLGEPPTAAVGPPTMASGGGAPMADNNEGA
DGVNASGNWHCDSTWLGDRTVITSTRTWALPTYNNHLYKQISSASTGASNDNHYFGY
STPWGYDFNRFCHFSPRDWORLINNNWGRPKRLNFKLFNQKVEVTTNDGVTTIA
NNTSTVQVFSDEYQLPYVLGSAHQCLPPFPADVPMIQQYGLTLNLSQAVGRSS
FYCLEYFPQMLRTGNFTPSYTFEDVPFHSYSAHSQSLDRLMNPLIDQYLYLNRTO
NQSQAQNKDLLFSRGPAGMSVQPKNWLPGCYQRQSVKTKTDNNNSNFTWTGASK
YNLNGRESIINPGTAMASHKDDKDKFFPMSGVMIFGKESAGASNTALDNVMTDEEEI
KATNPVATERFGTVAVNLQSSSTDPAVDVHMVGMALPGMVWQDRDVLQGPIMWAKIPH
TDGHEFSPLMGGFGLKHPPQILIKNTVPANPPAEFSATKFEASFITQYSTGQVSVE
IEWELOENSKRWNPVEQYTSNTAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL"
BASE COUNT 1114 a 1363 c 1277 g 929 t
ORIGIN
Alignment Scores:
Pred. No.: 2,04e-229 Length: 4683
Score: 3229.00 Matches: 594
Percent Similarity: 99.67% Conservative: 3
Best Local Similarity: 99.17% Mismatches: 2
Query Match: 99.32% Indels: 0
DB: 14 Gaps: 0
US-09-807-802A-15 (1-599) x AF028704 (1-4683)
Qy 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20
Db 2619 ACGGCTCCTGGAAAGAAACGTCCGGTAGAGCAGTCGCCACAAGAGCCAGACTCCTCCTCG 2678
Qy 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40
Db 2679 GGCATTGGCAAGACAGGCCAGCCCGCTAAAGAGAGACTCAATTTTGGTCAGACTGGC 2738
Qy 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60
Db 2739 GACTCAGAGTCAGTCCCGCAGCCCAACCTCTCGAGAGAACCTCCAGCAACCCCGCTGCT 2798
Qy 61 ValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGly 80
Db 2799 GTGGGACCTACTACATGGCTTCAGGCGGTGGCGCACCACCAATGGCAGACAATAACGAAGC 2858
Qy 81 AlaAspGlyValGlyAsnAlaSerGlyAsnThrHisCysAspSerThrTrpLeuGlyAsp 100
Db 2859 GCCGACGGAGTGGGTAAATGCCTCAGGAAATTTGGCATTCGATTCCACATGGCTGGCGAC 2918
Qy 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120
```



Db 2919 AGAGTCATCACCAACAGCACCGAACATGGCCCTTGCCACCTATAACAACCACTCTAC 2978  
Qy 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140  
Db 2979 AAGCAATCTCCAGTGTCTCAACGGGGGCCAGCAACGAAACCACTACTTCGGGTACAGC 3038  
Qy 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160  
Db 3039 ACCCCCTGGGGTATTTTGAATTTCAACAGATTCCACTGCCATTCTCACCAACGTGACTGG 3098  
Qy 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180  
Db 3099 CAGCGACTCATCAACAACAATTTGGGATTCGGGCCCAAGAGACTCAACTTCAAGCTCTTC 3158  
Qy 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200  
Db 3159 AACATCCAAGTCAAGGAGGTACGACGAATGATGGCGTCACGACCATCGCTAATAACCTT 3218  
Qy 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220  
Db 3219 ACCAGCAGGTTCAAGTCTTCTCGACTCGGAGTACCAGTTCGGGTACGTCTCGGCTCT 3278  
Qy 221 AlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGly 240  
Db 3279 GCGCACAGGGCTGCCTCCCTCCGTTCCCGCGGACGTGTTCATGATTCCGCGAGTACGGC 3338  
Qy 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260  
Db 3339 TACCTAACGCTCAACAATGGCAGCCAGGCAGTGGGACGTGCATCTCTTTACTGCGCTGAA 3398  
Qy 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280  
Db 3399 TATTTCCCATCGCAGATGCTGAGAACGGGCAATAACTTTACCTTCAGCTACACCTTCGAG 3458  
Qy 281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300  
Db 3459 GACGTGCCTTTCCACAGCAGCTACGCGCACAGCCAGAGCCTGGACGGCTGATGAATCTT 3518  
Qy 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320  
Db 3519 CTCATCGACCACTACCTGTATTACCTGAACAGAACTCAGATCAGTCCGGAAGTGCCCAA 3578  
Qy 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340  
Db 3579 AACAAGGACTTGCTGTTTAGCCGGGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAAC 3638  
Qy 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360  
Db 3639 TGGCTACCTGGACCTGTTTACCGGCAGCAGCGCGTCTTCTAAACAAAAACAGACAACAAC 3698  
Qy 361 AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380  
Db 3699 AACAGCAACTTTACCTGGACTGGTGTCTTCAAAATATAACCTTAATGGCGTGAATCTATA 3758  
Qy 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400  
Db 3759 ATCAACCCCTGGCACTGCTATGGCCTCACACAAGACGACAAGACAAGTCTTTTCCCATG 3818  
Qy 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420  
Db 3819 AGCGGTGTCATGATTTTGGAAAGGAGAGCGCGGAGCTTCAAAACACTGCATTGGACAAT 3878  
Qy 421 ValMetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440  
Db 3879 GTCATGATCACAGCAAGAGGAAATCAAAGCCACTAAACCCCGTGGCCACCGAAAGATT 3938  
Qy 441 GlyThrValAlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHis 460  
Db 3939 GGGACTGTGGCAGTCAATCTCCAGAGCAGCAGCAGACAGACCTTCGACCGGAGATGTGCAT 3998  
Qy 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 480

Db 3999 GTTATGGGAGCCTTACCTGGAATGGTGTGGCAAGACAGAGACGTATACCTGCAGGGTCCT 4058  
Qy 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500  
Db 4059 ATTTGGGCCAAAATTCCTCACACGGATGGACACTTTCACCCGCTCTCCTCTCATGGCGGC 4118  
Qy 501 PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520  
Db 4119 TTTGGACTTAAGCACCCGCTCTCTCAGATCCTCATCAAAAACACGCTGTCTCTGCGAAT 4178  
Qy 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540  
Db 4179 CCTCCGGCAGAGTTTTCGGCTACAAAGTTTGCTTCATTCATCACCAGTATTCACAGGA 4238  
Qy 541 GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro 560  
Db 4239 CAAGTGAGCGTGGAGATTGAATGGGAGCTGCAGAAAGAAACAGCAACCGCTGGAATCCC 4298  
Qy 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580  
Db 4299 GAAGTGCAGTATACATCTAACTATGCAAAATCTGCCAACGTTGATTTCTACTGTGGACAAC 4358  
Qy 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599  
Db 4359 AATGGACTTTATACTGAGCCTCGCCCCATTTGGCACCCTGTACCTCACCCGTCCTCG 4415  
RESULT 3  
AF028705 4722 bp DNA linear VRL 12-JAN-1998  
LOCUS Adeno-associated virus 3B, complete genome.  
DEFINITION AF028705  
ACCESSION AF028705  
VERSION AF028705.1 GI:2766608  
KEYWORDS  
SOURCE Adeno-associated virus 3B  
ORGANISM Adeno-associated virus 3B  
REFERENCE 1 (bases 1 to 4722)  
AUTHORS Rutledge,E.A., Halbert,C.L. and Russell,D.W.  
TITLE Infectious clones and vectors derived from adeno-associated virus (AAV) serotypes other than AAV type 2  
JOURNAL J. Virol. 72 (1), 309-319. (1998)  
MEDLINE 98080418  
PUBMED 9420229  
REFERENCE 2 (bases 1 to 4722)  
AUTHORS Rutledge,E.A. and Russell,D.W.  
TITLE Direct Submission  
JOURNAL Submitted (25-SEP-1997) Hematology, Univ of Washington, Box 357720, Seattle, WA 98195, USA  
FEATURES Location/Qualifiers  
source 1..4722  
/organism="Adeno-associated virus 3B"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:68742"  
/note="AAV3B"  
317..2191  
/codon\_start=1  
/product="nonstructural protein Rep78"  
/protein\_id="AAB95451.1"  
/db\_xref="GI:2766609"  
/translation="MPGFYEIVLKVPSDLDEHLPGISNPFVNWVAEKELPPDSMD  
PNLIEAPLTVAEKLQREFLVEWRRVSKAPEALFFVQFEKGETYFHLHLIETIGVKS  
MVVGRYVSQIKEKLVTRIYRGVEPQLPNWFAVTKTRNGAGGKNVDDCYIPNYLLPK  
TQPELQWANTNMDQYLSACLNLAEKRLVAQHLTHVSQTQEQNKENQNPNSDAPVIRS  
KTSARYMELVGLVDRGITSEKOWIQEDQASYISFNAAASNSRSQIKAAALDNASKIMSL  
TKTAPDYLVSNPEDITKNRIYQILELNGYDPPQYAAASVFLGWAQKKFGKRNITWLPF  
PATTKTNIAEALAHAVPFYGCVTWNTNENFPNDCVDKMWIWEEGKMTAKVVESAKA  
ILGSKVRVDQCKSSAQIEPTPIVTSNTNMCAVIDGNSSTTHQQLQDRMPKPEL  
TRRLDHFQKVTQEVKDFFRWASDHVTVHAHFVVRKGAKRPPASNDADVSEPKRQ  
CTSLAQPTTSDAAPADYADRYQNKCSRHVGMNMLPFPCKTCERMNQISNVCTHGQR  
DCGECFPGMSESPVSVVKKTYQKLCPIHILGRAPEIACSCDLANVLDLDCVSEQ  
"  
2208..4418  
/codon\_start=1  
CDS

```
/product="capsid protein vpi"
/protein_id="AAB95452.1"
/db_xref="GI:2766610"
/translation="MAADGYLPDWLEDNLSEIREWALKPGVPOPKANQQHQDNRRL
LVLPGYKYLPGNGLDKGEFVNEADAALEHDKAYDQQLKAGDNPYLKYNHADAEPQE
RLQEDTSFGGNLGRAVFAKKRILEPLGLVEEAAKTA PGKKRPVDQSPQEPDSSSGVG
KSGKQPARKLNFGQTGDSVPDPQPLGEPFAAPTSLGSNTWASGGGAPMADNNEGA
DGVNSSGNWHDSDQWLGDRIITSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYS
TPWGYFDNFRFHCHFSPRDWORLLINNWGFRPKLSFKLFNIQVKEVTQNDGTTTTIAN
NLTSVQVFTDSEYQLPYVLGSAHQCLPPFPADVMVPQYGLTLNNGSQAVGRSSF
YCLEYFSPQMLRTGNPFQSYTFEDVPFHSSYAHQSQSLDRLMNPLIDQYLYLNRTOG
TTSGTTNQSRLLFSQAGPQSMQLQARNWLPGEFCYRQRLSKTANDNNNSNFPWTAASK
YHLNGRDSLVPNPGPAMASHKDEEKFFPMHGNLIFGKEGTTASNAELDNVMTDEEEI
RTTNPVATEQYGTVANNLQSSNTAPTTRTVNDQGALPGMWVQDRDYLQGPWAKIPH
TDGHEHPSPLMGGFGLKHPPPIQIMIKNTPVPANPPTTFESPAKFASFITQYSTGQVSVE
IEWELQKENSKRWNPEIQYTSNYSKSVNVDFTVDTNGVYSEPRPIGTRYLTRLN"
BASE COUNT 1204 a 1297 c 1243 g 978 t
ORIGIN
```

```
Alignment Scores:
Pred. No.: 5.47e-200 Length: 4722
Score: 2832.00 Matches: 514
Percent Similarity: 91.33% Conservative: 34
Best Local Similarity: 85.67% Mismatches: 50
Query Match: 87.11% Indels: 2
DB: 14 Gaps: 2
```

US-09-807-802A-15 (1-599) x AF028705 (1-4722)

|    |      |  |      |
|----|------|--|------|
| QY | 1    | ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer   | 20   |
| Db | 2619 | ACGGCTCTGGAAAGAGAGCGCTGTAGATCAGTCTCTCAGGAACCGGACTCATCT         | 2678 |
| QY | 21   | GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly   | 40   |
| Db | 2679 | GGTGTGGCAAAATCGGGCAACAGCCCTGCCAGAAAAGACTAAATTTTCGGTCAGACTGGC   | 2738 |
| QY | 41   | AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla   | 60   |
| Db | 2739 | GACTCAGAGTCAGTCCCAGACCTCAACCTCTCGGAGAACCAACAGCAGCCCCCAAGT      | 2798 |
| QY | 61   | ValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGly      | 80   |
| Db | 2799 | TTGGGATCTAATACAAATGGCTTCAGCGGTGGCGCACCAATGGCAGACAATAACGAGGGT   | 2858 |
| QY | 81   | AlaAspGlyValGlyAsnAlaSerGlyAsnTirHisCysAspSerThrTirPleuGlyAsp  | 100  |
| Db | 2859 | GCCGATGGAGTGGTAAATTCCTCAGGAAATTTGGCATTGCGATTCCCAATGGCTGGCGAC   | 2918 |
| QY | 101  | ArgValIleThrThrSerThrArgThrTirPAlaLeuProThrTyrAsnAsnHisLeuTyr  | 120  |
| Db | 2919 | AGAGTCATCACCCAGCACCAACCTGGGGCCCTGCCCATTTACAACAACCATCTCTAC      | 2978 |
| QY | 121  | LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer   | 140  |
| Db | 2979 | AAGCAAATCTCCAGCCAATCA--GGAGCTTCAAACGACAACCACTACTTTGGCTACAGC    | 3035 |
| QY | 141  | ThrProTirPglyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTirP | 160  |
| Db | 3036 | ACCCCTTGGGGGTATTTTGACTTTAAACAGATTCCACTGCCACTTCTCACCACCGTACTGG  | 3095 |
| QY | 161  | GlnArgLeuIleAsnAsnAsnTirPglyPheArgProLysArgLeuAsnPheLysLeuPhe  | 180  |
| Db | 3096 | CAGCGACTCATTAACAACAACCTGGGGATTCCGGCCCCAAGAAACTCAGCTTCAAGCTCTTC | 3155 |
| QY | 181  | AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu   | 200  |
| Db | 3156 | AACATCCAAGTTAAAGAGGTACCGCAGAACCGATGGCAGCAGCACTATTGCCAATAACCTT  | 3215 |
| QY | 201  | ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer   | 220  |
| Db | 3216 | ACCAGCACGGTTCAAGTGTTTACGGACTCGGAGTATCAGCTCCCGTACGTCTCGGGTCG    | 3275 |

|    |      |   |      |
|----|------|---|------|
| QY | 221  | AlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGly      | 240  |
| Db | 3276 | GCGCACCAAGCTGTCTCCCGCCGTTTCCAGCGGACGTCTTTCATGGTCCCTCAGTATGGA      | 3335 |
| QY | 241  | TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu      | 260  |
| Db | 3336 | TACCTCACCTGAAACAACGGAAGTCAAGCGGTGGACGCTCATCTTTTACTGCCTGGAG        | 3395 |
| QY | 261  | TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu      | 280  |
| Db | 3396 | TACTTCCCTTCGCAGATGCTAAGGACTGGAATAAATTCCTTTCAGCTATACCTTCGAG        | 3455 |
| QY | 281  | GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro      | 300  |
| Db | 3456 | GATGTACCTTTTCACAGCAGCTACGCTCACAGCCAGAGTTTGGATCGCTTGATGAATCCT      | 3515 |
| QY | 301  | LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGln---AsnGlnSerGlySerAla      | 319  |
| Db | 3516 | CTTATTGATCAGTATCTGTACTACCTGAACAGAACGCAAGGAACAACCTCTGGAACAACC      | 3575 |
| QY | 320  | GlnAsnLysAspLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLys         | 339  |
| Db | 3576 | AACCAATCACGGCTGCTTTTAGCCAGCTGGGCCTCAGTCTATGTCTTTGCAGGCCAGA        | 3635 |
| QY | 340  | AsnTirPleuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsn     | 359  |
| Db | 3636 | AATTGGCTACTGGGCCCTGTCTACCGCAACAGAGACTTTCAAAGACTGCTAACGACAAC       | 3695 |
| QY | 360  | AsnAsnSerAsnPheThrTirPThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSer     | 379  |
| Db | 3696 | AACAACAGTAACTTCTCTGGACAGCGGCAGCAAAATATCATCTCAATGGCCGCGACTCG       | 3755 |
| QY | 380  | IleIleAsnProGlyThrAlaMetAlaSerHisLysAspGluAspLysPhePhePro         | 399  |
| Db | 3756 | CTGGTGAATCCAGGACCAGCTATGGCCAGTCACAGGACGATGAAGAAAAATTTTCCCT        | 3815 |
| QY | 400  | MetSerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAsp      | 419  |
| Db | 3816 | ATGCACGGCAATCTAATATTTGGCAAAAGAGGACAACGGCAAGTAACGCAGAAATTAGAT      | 3875 |
| QY | 420  | AsnValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArg         | 439  |
| Db | 3876 | AATGTAATGATTACGGATGAAGAAGAGATTTCGTACCACCAATCCTGTGGCAACAGAGCAG     | 3935 |
| QY | 440  | PheGlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspVal      | 459  |
| Db | 3936 | TATGGAACCTGGCAAAATAACTTTCAGAGACTCAAATACAGCTCCCACGACTAGAACTGTC     | 3995 |
| QY | 460  | HisAlaMetGlyAlaLeuProGlyMetValTirPTrpGlnAspArgAspValTyrLeuGlnGly  | 479  |
| Db | 3996 | AATGATCAGGGGCGCTTACCTGGCATGTGTGGCAAGATCGTGACGTGCTGCTTCAAGGA       | 4055 |
| QY | 480  | ProIleTirPAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGly     | 499  |
| Db | 4056 | CCTATCTGGCAAAAGATTCTCTCACACGGATGGACACTTTCATCCTTCTCTCTCTGATGGGA    | 4115 |
| QY | 500  | GlyPheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAla         | 519  |
| Db | 4116 | GGCTTTGGACTGAAACATCCGCGCTCTCAAATCATGATCAAAAATACTCCGGTACCGGCA      | 4175 |
| QY | 520  | AsnProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThr      | 539  |
| Db | 4176 | AATCCTCCGACGACTTTCAGCCCCGGCCCAAGTTTGCTTCATTTATCACTCAGTACTCCACT    | 4235 |
| QY | 540  | GlyGlnValSerValGluIleGluTirPTrpGluLeuGlnLysGluAsnSerLysArgTirPAsn | 559  |
| Db | 4236 | GGACAGGTCAGCGTGGAAATTGAGTGGGAGCTACAGAAAGAAAACAGCAACGTTTGGAAAT     | 4295 |
| QY | 560  | ProGluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAsp      | 579  |
| Db | 4296 | CCAGAGATTTCAGTACACTTCCCAACTACAAACAAGTCTGTATATGTGGACTTTACTGTAGAC   | 4355 |
| QY | 580  | AsnAsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu      | 599  |





Db 3577 AACCAATCACGGCTGCTTTTAGCCAGGCTGGCCCTCAGTCTATGTCTTTGCAGGCCAGA 3636

Qy 340 AsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsn 359

Db 3637 AATTGGTACCTGGGCCCTGCTACCGGCAACAGAGACTTTCAAAGACTGCTTAACGACAAC 3696

Qy 360 AsnAsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSer 379

Db 3697 AACCAACAGTAACCTTTCCTTGGACAGCGGCCAGCAATATCATCTCAATGGCCGCACTCG 3756

Qy 380 IleIleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePhePro 399

Db 3757 CTGGTGAATCCAGGACCAGCTATGGCCAGTCACAGGACGATGAAGAAAAATTTTTCCT 3816

Qy 400 MetSerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAsp 419

Db 3817 ATGCACGGCAATCTAATAATTTGGCAAGAAGGACACGGCAAGTAACGCAGATTAGAT 3876

Qy 420 AsnValMetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArg 439

Db 3877 AATGTAATGATTACGGATGAAGAAGAGATTTCGTACCAACAATCCTGTGGCAACAGAGCAG 3936

Qy 440 PheGlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspVal 459

Db 3937 TATGGAACTGTGGCAAAATACTGCAGAGCTCAAATACAGTCCCACGACTGGAACTGTC 3996

Qy 460 HisAlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGly 479

Db 3997 AATCATCAGGGGGCCTTACCTGGCATGGTGTGGCAAGATCGTGACGTGTACCTTCAAGGA 4056

Qy 480 ProlleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGly 499

Db 4057 CCTATCTGGGCAAGATTCTTCACACGGATGGACACTTTCATCTTCTCCTCTGATGGGA 4116

Qy 500 GlyPheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAla 519

Db 4117 GGCTTTGGACTGAACATCCGCCTCCTCAATCATGATCAAAAATACTCCGGTACCGCA 4176

Qy 520 AsnProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThr 539

Db 4177 AATCTCCGACGACTTTCAGCCCGGCCAAGTTTGCTTTCATTTATCACTCAGTACTCCACT 4236

Qy 540 GlyGlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsn 559

Db 4237 GGACAGGTCAGCGTGGAAATTGAGTGGGAGCTACAGAAAGAAAACAGCAACGTTTGAAT 4296

Qy 560 ProGluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAsp 579

Db 4297 CCAGAGATTCACTACACTTCCAACTACAACAAGTCTGTTAATGTGGACTTTACTGTAGAC 4356

Qy 580 AsnAsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599

Db 4357 ACTAATGGTGTATTATAGTGAACCTCGCCCTATTGGAACCCGGTATCTCACACGAAACTTG 4416

RESULT 5

AR264580

LOCUS AR264580 8179 bp DNA linear PAT 10-APR-2003

DEFINITION Sequence 5 from patent US 6491907.

ACCESSION AR264580

VERSION AR264580.1 GI:29692850

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 8179)

AUTHORS Rabinowitz,J.E., Samulski,R.J. and Xiao,W.

TITLE Recombinant parvovirus vectors and method of making

JOURNAL Patent: US 6491907-A 5 10-DEC-2002;

FEATURES Location/Qualifiers

source 1..8179

/organism="unknown"

BASE COUNT 2082 a 2078 c 2027 g 1992 t

ORIGIN

Alignment Scores: 2.28e-197 Length: 8179

Pred. No.: 2801.00 Matches: 509

Score: 90.33% Conservative: 33

Percent Similarity: 84.83% Mismatches: 56

Best Local Similarity: 86.16% Indels: 2

Query Match: Gaps: 2

DB: 6

US-09-807-802A-15 (1-599) x AR264580 (1-8179)

Qy 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20

Db 2544 ACGGCTCCGGGAAAAAAGAGGGCCGGTAGAGCACTCTCCTGTGGAGCCAGACTCCTCCTCG 2603

Qy 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40

Db 2604 GGAACCGGAAAGCGGGCCAGCAGCCTTGCAGAAAAGATTGAATTTGGTCAGACTGGA 2663

Qy 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60

Db 2664 GACGCAGACTCAGTACCTGACCCCCAGCCTCTCGGACAGCCACCAGACGCCCTCTTGGT 2723

Qy 61 ValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGly 80

Db 2724 CTGGGAACCTAATACGATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGGC 2783

Qy 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100

Db 2784 GCCGACGGAGTGGGTAATCTCTCCGAAATTTGGCATTCGATTCCACATGATGGGCGAC 2843

Qy 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120

Db 2844 AGAGTCATCACCAACAGCACCCGAACTGGGCCCTGCCACCTACAACAACACCTCTAC 2903

Qy 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140

Db 2904 AAACAAATTTCCAGCCCAATCA--GGAGCCCTCGAACGACAATCACTACTTTGGCTACAGC 2960

Qy 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160

Db 2961 ACCCTTGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCCACCGTACTGG 3020

Qy 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180

Db 3021 CAAAGACTCATCAACAACAACCTGGGATTCGACCCCAAGAGACTCAACTTCAAGCTCTTT 3080

Qy 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200

Db 3081 AACATTCAAGTCAAAGAGGTCAACGAGATGACGATGACGACGATTCGCAATAACCTT 3140

Qy 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220

Db 3141 ACCAGCAGGTTCAAGTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTCTCGGGTCTG 3200

Qy 221 AlaHisGlnGlyCysLeuProPheProPheProAlaAspValPheMetIleProGlnTyrGly 240

Db 3201 GCGCACCAAGGCTGTCTCCCGCCGTTTCCAGCGGACGTCTTTCATGGTCCCTCAGTATGGA 3260

Qy 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260

Db 3261 TACCTCACCCCTGAACAACGGAAGTCAAGCGGTGGGACGCTCATCTTTTACTGCTGGAG 3320

Qy 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280

Db 3321 TACTTCCCTTCGCAGATGCTAAGGACTGGAATAAATACTTCCAATTCACTATACCTTCGAG 3380

Qy 281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300

Db 3381 GATGTACCTTTTCACAGCAGCTACGCTCACAGCCAGAGTTTGGATCGCTTGATGAATCCT 3440

Qy 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGln---AsnGlnSerGlySerAla 319



|            |                                   |   |                            |
|------------|-----------------------------------|---|----------------------------|
| Db         | 2732                              | GACGTTCTCTTCCACAGCAGCTACGCTCACAGCCACAGTCTGGACCGTCTCATGAATCCT    | 2791                       |
| Qy         | 301                               | LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln    | 320                        |
| Db         | 2792                              | CTCATCGACCAGTACTGTATTACTTTGAGCAGAAACAACACTCCAAGTGGAAACCACG      | 2851                       |
| Qy         | 321                               | AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn    | 340                        |
| Db         | 2852                              | CAGTCAAGGCTTCAGTTTTTCTCAGGCCGGAGCGAGTGACATTCCGGACCAGTCTAGGAAC   | 2911                       |
| Qy         | 341                               | TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn    | 360                        |
| Db         | 2912                              | TGGCTTCCTGGACCCCTGTTACCGCCAGCAGCGAGTATCAAAAGACATCTGCGGATAACAAC  | 2971                       |
| Qy         | 361                               | AanSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle    | 380                        |
| Db         | 2972                              | AACAGTGAATACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTG    | 3031                       |
| Qy         | 381                               | IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet    | 400                        |
| Db         | 3032                              | GTGAATCCGGGCGCCGTCATCTTTGGGAAGCAAGGCTCAGAGAAACAATGTGGACATTGAAAG | 3091                       |
| Qy         | 401                               | SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn    | 420                        |
| Db         | 3092                              | AGCGGGTTCTCATCTTTGGGAAGCAAGGCTCAGAGAAACAATGTGGACATTGAAAG        | 3151                       |
| Qy         | 421                               | ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe       | 440                        |
| Db         | 3152                              | GTTCATGATTACAGACGAAGAGAAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTAT    | 3211                       |
| Qy         | 441                               | GlyThrValAlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHis       | 460                        |
| Db         | 3212                              | GGTCTGTATCTACCAACCTCCAGAGAGGCAACAGACAAGCAGCTACCCGAGATGTCAAC     | 3271                       |
| Qy         | 461                               | AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro    | 480                        |
| Db         | 3272                              | ACACAAGCGGTTCTTCCAGGCATGGTCTGGCAGGACAGAGATGTACCTTCAGGGGCC       | 3331                       |
| Qy         | 481                               | IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly    | 500                        |
| Db         | 3332                              | ATCTGGGCAAGATTCCACACACGACGGACGGACATTTTCAACCCCTCTCCCTCATGGGTGA   | 3391                       |
| Qy         | 501                               | PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn       | 520                        |
| Db         | 3392                              | TTCCGACTTAAACACCCCTCCTCCACAGATTCTCATCAAGAACACCCCGGTACCTCGGAAT   | 3451                       |
| Qy         | 521                               | ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly    | 540                        |
| Db         | 3452                              | CCTTCGACCACCTTCAGTGGCGCAAAGTTTGCTTCTTCATCACACAGTACTCCACGGGA     | 3511                       |
| Qy         | 541                               | GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro    | 560                        |
| Db         | 3512                              | CAGTTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAAACAGCAACGCTGGAATCCC      | 3571                       |
| Qy         | 561                               | GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn    | 580                        |
| Db         | 3572                              | GAAATTTCAGTACACTTCCAACTACAACAAGTCTGTTAATGTGGACTTTACTGTGGACACT   | 3631                       |
| Qy         | 581                               | AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu       | 599                        |
| Db         | 3632                              | AATGGCGTGATTTCAGAGCCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTG       | 3688                       |
| RESULT 7   |                                   |   |                            |
| AX205074   |                                   |   |                            |
| LOCUS      | AX205074                          | 4072 bp   | DNA linear PAT 30-AUG-2001 |
| DEFINITION | Sequence 4 from Patent WO0155361. |   |                            |
| ACCESSION  | AX205074.                         |   |                            |
| VERSION    | AX205074.1 GI:15394315            |   |                            |
| KEYWORDS   | .                                 |   |                            |
| SOURCE     | unidentified                      |   |                            |
| ORGANISM   | unidentified                      |   |                            |
| REFERENCE  | 1                                 |   |                            |

|   |      |  |               |       |       |
|---|------|--|---------------|-------|-------|
| AUTHORS                                       |      | Hardy, S.F.  |               |       |       |
| TITLE   |      | Recombinant aav packaging systems                              |               |       |       |
| JOURNAL                                       |      | Patent: WO 0155361-A 4 02-AUG-2001;                            |               |       |       |
| FEATURES                                      |      | Chiron Corporation (US)  |               |       |       |
|   |      | Location/Qualifiers  |               |       |       |
|   |      | 1..4072  |               |       |       |
|   |      | /organism="unidentified"                                       |               |       |       |
|   |      | /mol_type="genomic DNA"  |               |       |       |
|   |      | /db_xref="taxon:32644"   |               |       |       |
|   |      | /note="recombinant DNA"  |               |       |       |
| BASE COUNT                                    |      | 1073 a   | 1025 c        | 992 g | 982 t |
| ORIGIN  |      |  |               |       |       |
| Alignment Scores:                             |      |  |               |       |       |
| Pred. No.:                                    |      | 4.51e-195  | Length:       | 4072  |       |
| Score:  |      | 2764.50  | Matches:      | 496   |       |
| Percent Similarity:                           |      | 89.98%   | Conservative: | 43    |       |
| Best Local Similarity:                        |      | 82.80%   | Mismatches:   | 59    |       |
| Query Match:                                  |      | 85.04%   | Indels:       | 1     |       |
| DB:   |      | 6  | Gaps:         | 1     |       |
| US-09-807-802A-15 (1-599) x.AX205074 (1-4072) |      |  |               |       |       |
| QY  | 1    | ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer   | 20            |       |       |
| DB  | 1895 | ACGGCTCCGGGAAAAAAGAGGCCGTAGAGCACTCTCCTGTGGAGCCAGACTCCTCCTCG    | 1954          |       |       |
| QY  | 21   | GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly   | 40            |       |       |
| DB  | 1955 | GGAACCGGAAAGCGGGCCAGAGCCTGCAAGAAAAAGATTGAATTTGGTCAGACTGGA      | 2014          |       |       |
| QY  | 41   | AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla   | 60            |       |       |
| DB  | 2015 | GACGCAGACTCAGTACCTGACCCCCAGCCTCTCGGACAGCCACCAGCCCCCTCTGGT      | 2074          |       |       |
| QY  | 61   | ValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGly   | 80            |       |       |
| DB  | 2075 | CTGGGAACCTAATACGATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGC   | 2134          |       |       |
| QY  | 81   | AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp   | 100           |       |       |
| DB  | 2135 | GCCGACGGAGTGGGTAATTCCTCGGGAATTTGGCATTTCCACATGGATGGGCGAC        | 2194          |       |       |
| QY  | 101  | ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr   | 120           |       |       |
| DB  | 2195 | AGAGTCATCACCCACAGCACCCGAACCTGGGCCCTGCCACCTACACAACACCTCTAC      | 2254          |       |       |
| QY  | 121  | LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer   | 140           |       |       |
| DB  | 2255 | AAACAAATTTCCAGCCAATCA--GGAGCCTCGAACGACAATCACTACTTGGCTACAGC     | 2311          |       |       |
| QY  | 141  | ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp   | 160           |       |       |
| DB  | 2312 | ACCCCTTGGGGGTATTTTGACTTTCACACAGATTCCACTGCCACTTTTACCACGTGACTGG  | 2371          |       |       |
| QY  | 161  | GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe   | 180           |       |       |
| DB  | 2372 | CAAAGACTCATCAACAACAACCTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTT | 2431          |       |       |
| QY  | 181  | AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu   | 200           |       |       |
| DB  | 2432 | AACATTTCAAGTCAAAGAGGTTCACGAGAATGACGGTACGACGATTGCCAATAACCTT     | 2491          |       |       |
| QY  | 201  | ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer   | 220           |       |       |
| DB  | 2492 | ACCAGCACGGTTCAGGTGTTTACTGACTCGGAGTACCAGCTCCCCGTACGTCTCTCGGCTCG | 2551          |       |       |
| QY  | 221  | AlaHisGlnGlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGly      | 240           |       |       |
| DB  | 2552 | GCGCATCAAGGATGCCCTCCCGCGTTCCAGCAGACGCTCTTTCATGGTGCCACAGTATGGA  | 2611          |       |       |
| QY  | 241  | TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu   | 260           |       |       |



Db 2612 TACCTCACCTGAACAACGGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCTGGAG 2671

Qy 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280  
|||||

Db 2672 TACTTTCCTTCTCAGATGCTGGTACCGGAACAACATTACCTTCAGCTACACTTTTGAG 2731

Qy 281 GluValPropheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300  
:::|||||

Db 2732 GACGTCCTTTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCT 2791

Qy 301 LeuLeuAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320  
|||||

Db 2792 CTCATCGACCAGTACCTGTATTACTTGAGCAGAACAAACACTCCAAGTGGAAACCAACG 2851

Qy 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340  
|||||

Db 2852 CAGTCAAGGCTTCAGTTTTCTCAGGCCGGAGCGAGTGACATTCGGGACCAGTCTAGGAAC 2911

Qy 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360  
|||||

Db 2912 TGGCTTCTTGACCCCTGTTACCGCCAGCAGCGAGTATCAAGACATCTCGGATATAACAAC 2971

Qy 361 AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380  
|||||

Db 2972 AACAGTGAATACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTG 3031

Qy 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400  
:::|||||

Db 3032 GTGAATCCGGGCCCGCCCATGGCAAGCCACAGGACGATGAAGAAAGTTTTTTCCTCAG 3091

Qy 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420  
|||||

Db 3092 AGCGGGGTCTCATCTTTGGGAAGCAAGGCTCAGAGAAACAAATGTGGACATTGAAGAAG 3151

Qy 421 ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440  
|||||

Db 3152 GTCATGATTACAGACGAAGAGGAAATCAGGACAAACCAATCCCGTGGCTACGGAGCAGTAT 3211

Qy 441 GlyThrValAlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHis 460  
|||||

Db 3212 GGTTCGTATCTACCAACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAAC 3271

Qy 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 480  
|||||

Db 3272 ACACAAGGCGTCTTCCAGGCATGGTCTGCAGGACAGAGATGTGTACCTTCAGGGGCC 3331

Qy 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500  
|||||

Db 3332 ATCTGGGCAAGATTCCACACGAGCGGACATTTTCAACCCCTCTCCCTCATGGGTGGA 3391

Qy 501 PheGlyLeuLysAsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520  
|||||

Db 3392 TTCGGACTTAAACACCCCTCCTCCACAGATTCTCATCAAGAAACACCCCGGTACCTGCGAAT 3451

Qy 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540  
|||||

Db 3452 CCTTCGACCACCTTCAGTGGCGGCAAGTTTGCTTCTCTCATCACAGTACTCCACGGGA 3511

Qy 541 GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro 560  
|||||

Db 3512 CAGGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAAACAGCAACCGTGGAAATCCC 3571

Qy 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580  
|||||

Db 3572 GAAATTCAGTACACTTCCAACACTACAACAAGTCTGTTAATGTGGACTTTACTGTGGACACT 3631

Qy 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599  
|||||

Db 3632 AATGGCGTGATTACAGAGCCTCGCCCCCATTTGGCACCAGATACCTTGACTCGTAATCTG 3688

RESULT 8  
AX282480  
LOCUS

DEFINITION Sequence 1 from Patent WO0168888.  
ACCESSION AX282480  
VERSION AX282480.1 GI:16609611  
KEYWORDS  
SOURCE Adeno-associated virus 2  
ORGANISM Adeno-associated virus 2  
REFERENCE Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.  
AUTHORS Xiao, W. and Daring, M.J.  
TITLE Production of chimeric capsid vectors  
JOURNAL Patent: WO 0168888-A 1 20-SEP-2001;  
Neurologix, Inc. (US)  
FEATURES Location/Qualifiers  
source  
1. 4679  
/organism="Adeno-associated virus 2"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10804"  
BASE COUNT 1198 a 1262 c 1255 g 964 t  
ORIGIN  
Alignment Scores: 5.44e-195 Length: 4679  
Pred. No.: 2764.50 Matches: 496  
Score: 89.98% Conservative: 43  
Percent Similarity: 82.80% Mismatches: 59  
Best Local Similarity: 85.04% Indels: 1  
Query Match: 6 Gaps: 1  
DB: 1  
US-09-807-802A-15 (1-599) x AX282480 (1-4679)  
Qy 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSer 20  
Db 2614 ACGGCTCCGGGAAAAAAGAGGCCGGTAGAGCACTCTCCTGTGGAGCCAGACTCCTCTCG 2673  
Qy 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40  
Db 2674 GGAACCGAAAGGGCGGCCAGCAGCCTGCAAGAAAAAGATTGAATTTTGGTCAGACTGGA 2733  
Qy 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60  
Db 2734 GAGCAGACTCAGTACTACCTGACCCCGAGCCTCTCGGACAGCCACCAGCAGCCCTCTG 2793  
Qy 61 ValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGly 80  
Db 2794 CTGGAACTAATACGATGGCTACAGGCAGTGGCGGACCAATGGCATTCGATTCACATGGATGGCGAC 2853  
Qy 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100  
Db 2854 GCCACGGAGTGGTAATTCTCGGAAATTGGCAATTCGATTCGATTCACATGGATGGCGAC 2913  
Qy 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120  
Db 2914 AGAGTCATCACCACCAGCACCAGAACCTGGGCCCTGCCCACCTACAACAACCACTCTAC 2973  
Qy 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140  
Db 2974 AAACAAATTTCCAGCCCAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGC 3030  
Qy 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160  
Db 3031 ACCCTTGGGGGTATTTTGACTTCAACAGATTCCACTGCCCACCTTTTCCACCACGTGACTGG 3090  
Qy 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180  
Db 3091 CAAAGACTCATCAACAACAACACTGGGATTCGACCCCAAGAGACTCAACTTCAAGCTCTTT 3150  
Qy 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200  
Db 3151 AACATTCAAGTCAAAGAGGTCAACGAGAATGACGGTACGCGTACGACGACGATTGCCAATAACCTT 3210  
Qy 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220  
Db 3211 ACCAGCAGGTTTCAGGTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTCCTCGGCTCG 3270

QY 221 AlaHisGlnGlyCysLeuProPropheProAlaAspValPheMetIleProGlnTyrGly 240  
Db GCGCATCAAGGATGCTCCGCGCGTTCCAGCAGACGCTCTTCATGGTGCACAGTATGGA 3330

QY 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260  
Db TACCTCACCTTGAACAACGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCCTGGAG 3390

QY 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280  
Db TACTTTCCTTCTCAGATGCTGGTACCGGAACAACCTTACCTTCAGCTACACTTTTGAG 3450

QY 281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300  
Db GACGTTCTCTTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCT 3510

QY 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320  
Db CTCATCGACCAGTACCTGTATTACTTGAGCAGAACAAACACTCCAAGTGGAAACCACG 3570

QY 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340  
Db CAGTCAAGGCTTCAGTTTCTCAGGCCGAGCGAGTGACATTCGGGACCACTAGGAAC 3630

QY 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360  
Db TGGCTTCTGGACCCCTGTACCGCCAGCAGCGAGTATCAAGACATCTCGCGATAACAAC 3690

QY 361 AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380  
Db AACAGTGAATACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTG 3750

QY 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400  
Db GTGAATCCGGGCCCGGCCATGCGCAAGCCACAGGACGATGAAGAAAGTTTTTCTCTCAG 3810

QY 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420  
Db AGCGGGTTCTCATCTTTGGGAAGCAAGGCTCAGAGAAACAAATGTGGACATTGAAAG 3870

QY 421 ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440  
Db GTCATGATTACAGACGAAGAGGAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTAT 3930

QY 441 GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis 460  
Db GGTTCTGTATCTACCAACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAAC 3990

QY 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 480  
Db ACACAAGCGGTTCTTCCAGGCGATGGTCTGGCAGGACAGAGATGTACCTTCAGGGGCC 4050

QY 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500  
Db ATCTGGGCAAAAGATTCCACACGAGCGGACATTTTCACCCCTCTCCCTCATGGGTGGA 4110

QY 501 PheGlyLeuLysAsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520  
Db TTCGGACTTAAACACCCCTCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAAT 4170

QY 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540  
Db CCTTCGACCACCTTCAGTGGGCAAAAGTTTGCTTCTTCATCACACAGTACTCCACGGGA 4230

QY 541 GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTyrPheAsnPro 560  
Db CAGGTACGGTGGAGATCGAGTGGAGCTGCAGAAGGAARAAACAGCAACCGCTGGATCCC 4290

QY 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580  
Db GAAATTCAGTACACTTCCAACTACAACAAGTCTGTTAATGTGGACTTTACTGTGGAACT 4350

QY 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599  
Db AATGGCGTGTATTTCAGAGCCTCGCCCCATTGGCACCAGATACCTGACTCGTAATCTG 4407

RESULT 9  
AF043303 4679 bp DNA linear VRL 24-FEB-1998  
LOCUS Adeno-associated virus 2, complete genome.  
DEFINITION AF043303  
ACCESSION AF043303.1 GI:2906016  
VERSION  
KEYWORDS  
SOURCE Adeno-associated virus 2  
ORGANISM Adeno-associated virus 2  
VIRUSES; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.  
REFERENCE 1 (bases 1 to 4679)  
AUTHORS Ruffing,M., Heid,H. and Kleinschmidt,J.A.  
TITLE Mutations in the carboxy terminus of adeno-associated virus 2 capsid proteins affect viral infectivity: lack of an RGD integrin-binding motif  
JOURNAL J. Gen. Virol. 75 (Pt 12), 3385-3392 (1994)  
MEDLINE 95088582  
PUBMED 7996133  
REFERENCE 2 (bases 1 to 4679)  
AUTHORS Berns,K.I., Bohenzky,R.A., Cassinotti,P., Colvin,D., Donahue,B.A., Dull,T., Horer,M., Kleinschmidt,J.A., Ruffing,M., Snyder,R.O., Tratschin,J.-D. and Weitz,M.  
TITLE Direct Submission  
JOURNAL Submitted (15-JAN-1998) Cell Genesys Inc., 342 Lakeside Dr., Foster City, CA 94404, USA  
FEATURES  
source  
1. .4679  
/organism="Adeno-associated virus 2"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10804"  
/note="changes relative to the original sequence, GenBank Accession Number J01901, have been detected and verified by several different laboratories"  
repeat\_region  
1. .145  
/note="inverted terminal repeat"  
/rpt\_type=inverted  
misc\_feature  
42. .83  
/note="flip oriented DNA"  
precursor\_RNA  
287. .4451  
CDS  
join(321. .1906,2228. .2252)  
/codon\_start=1  
/product="Rep 68 protein"  
/protein\_id="AAC03774.1"  
/db\_xref="GI:2906017"  
/translation="MPGFYEIVIKVPSDLDEHLPGISDSFVNWVAEKELPPDSMD  
LNLIQAPLTVAEKLQDFLTWRVRVSKAPEALFFVQFEKGESYFHMVLTETTVGVS  
MVLGRFLSQIREKLIQRIYRGIEPTLPNWFVAVTKTRNGAGGNKVVDCEYIPNYLLPK  
TQPELQAWNTNMEQYLSACLNLTERKRLVAQHLTHVSQTQEQNKENQNSDAPVIRS  
KTSARYMELVGLVDKGITSEKQWIOEQDQASYISFNAASNSRSQIKAALDNAGKIMSL  
TKTAPDYLVGQPVEDISSNRIYKILELNGYDPQYAAVSFLGWATKKFGKRNITWLPFG  
PATTGKTNIABEIAHTVPFYGCYNWTNENFPFNDCCVDMVIWEEGKMTAKVVESAKA  
ILGGSKVRVDQCKSSAQIDPTPVIIVTSNTNMCVIDGNSTTFEHOQPLQDRMPKFEL  
TRRLDHFQKVTQEVKDFFRWAKDHVVEVEHEFYVKKGGAKKRPAPSDADISEPKRV  
RESVAQPSTSDAEASINYADRLARGHSL"  
321. .2186  
/codon\_start=1  
/product="Rep 78 protein"  
/protein\_id="AAC03775.1"  
/db\_xref="GI:2906018"  
/translation="MPGFYEIVIKVPSDLDEHLPGISDSFVNWVAEKELPPDSMD  
LNLIQAPLTVAEKLQDFLTWRVRVSKAPEALFFVQFEKGESYFHMVLTETTVGVS  
MVLGRFLSQIREKLIQRIYRGIEPTLPNWFVAVTKTRNGAGGNKVVDCEYIPNYLLPK  
TQPELQAWNTNMEQYLSACLNLTERKRLVAQHLTHVSQTQEQNKENQNSDAPVIRS  
KTSARYMELVGLVDKGITSEKQWIOEQDQASYISFNAASNSRSQIKAALDNAGKIMSL  
TKTAPDYLVGQPVEDISSNRIYKILELNGYDPQYAAVSFLGWATKKFGKRNITWLPFG  
PATTGKTNIABEIAHTVPFYGCYNWTNENFPFNDCCVDMVIWEEGKMTAKVVESAKA  
ILGGSKVRVDQCKSSAQIDPTPVIIVTSNTNMCVIDGNSTTFEHOQPLQDRMPKFEL  
TRRLDHFQKVTQEVKDFFRWAKDHVVEVEHEFYVKKGGAKKRPAPSDADISEPKRV  
RESVAQPSTSDAEASINYADRYQNKSRHVGMNLMFLPCRQCBRMNQNSNICFTTHGQK





Db 2614 ACGGCTCCGGGAAAAAAGAGGGCGGTAGAGCACTCTCCTGTGGAGCCAGACTCCTCCTCG 2673

Qy 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40

Db 2674 GGAACCGGAAGCGGGCCAGCAGCGCTGCAAGAAAAAGATTGAATTTTGGTCAGACTGGA 2733

Qy 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60

Db 2734 GACGCACTCAGTACCTGACCCCGCAGCTCTCGGACAGCCACAGCAGCCCCCTCTGGT 2793

Qy 61 ValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnGluGly 80

Db 2794 CTGGGAACATAACGATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGGC 2853

Qy 81 AlaAspGlyValGlyAsnAlaSerGlyValAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100

Db 2854 GCCACGGAGTGGTAATTCTCGGGAATGGCATTTGCGATTCCACATGGATGGGCGAC 2913

Qy 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120

Db 2914 AGAGTCATCACCAACAGCACCCGAACTTGGGCCCTGCCACCTACAAACACCACTCTAC 2973

Qy 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140

Db 2974 AAACAAATTTCCAGCCAATCA--GGAGCTCGAACGACAATCACTACTTTTGGCTACAGC 3030

Qy 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160

Db 3031 ACCCTTGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCCACCACGTGACTGG 3090

Qy 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180

Db 3091 CAAAGACTCATCAACAACAACCTGGGATTCGACCCCAAGAGACTCAACTTCAAGCTCTTT 3150

Qy 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200

Db 3151 AACATTCAAGTCAAAGAGGTCACGCAGAAATGACGGTACGACGACGATTGCCAATAACCTT 3210

Qy 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220

Db 3211 ACCAGCAGGTTTCAAGTGTACTGACTCGGAGTACCAGTCCCGTACGTCCTCGGCTCG 3270

Qy 221 AlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGly 240

Db 3271 GCGCATCAAGGATGCTTCCCGCTCCCGCTTCCAGCAGACGCTCTTCATGGTGCACAGTATGA 3330

Qy 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260

Db 3331 TACCTCACCTGAAACAACGGGAGTCAGGAGTAGGACGCTCTTCAATTTACTGCTGGAG 3390

Qy 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280

Db 3391 TACTTTCTCTCAGATGCTGCGTACCGGAAACAACCTTTACCTTACGCTACACTTTTGAG 3450

Qy 281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300

Db 3451 GACGTTCTCTTCCACAGCAGCTACGCTCAGCCAGAGTCTGGACCGTCTCATGAATCCT 3510

Qy 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320

Db 3511 CTCATCGACCACTACCTGTATTACTTTGAGCAGAAACAACACTCCAAGTGGAAACCCACG 3570

Qy 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340

Db 3571 CAGTCAAGGCTTCAGTTTCTCAGGCCGAGCGAGTGACATTCGGGACCAGTCTAGGAAC 3630

Qy 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360

Db 3631 TGGCTTCTCGGACCTGTTTACCGCCAGCAGCGAGTATCAAAGACATCTGCGGATAACAAC 3690

Qy 361 AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380

Db 3691 AACAGTGAATACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTG 3750

Qy 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400

Db 3751 GTGAATCCGGGCCCCGCGCATGGCAAGCCACAAGGACGATGAAGAAAGTTTTTTCCTCAG 3810

Qy 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420

Db 3811 AGCGGGTTCTCATCTTTGGGAAGCAAGGCTCAGAGAAAACAATGTGGACATTGAAAAG 3870

Qy 421 ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440

Db 3871 GTCATGATTACAGACGAAGAGGAATCAGGACAACCAATCCCCTGGCTACGGAGCAGTAT 3930

Qy 441 GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis 460

Db 3931 GGTTCGTATCTACCAACCTCCAGAGAGGCAACAGACAGCTACCGCAGATGTCAAC 3990

Qy 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 480

Db 3991 ACACAAGCGGTTCTTCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCC 4050

Qy 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500

Db 4051 ATCTGGGCAAGATTCCACACACGCGACGACATTTTCCACCCCTCTCCCTCATGGGTGA 4110

Qy 501 PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520

Db 4111 TTGGACTTAAACACCCCTCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAAT 4170

Qy 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540

Db 4171 CCTTCGACCACCTTCAGTGGGCAAGTTTGCTTCTTCATCACACAGTACTCCACGGGA 4230

Qy 541 GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro 560

Db 4231 CAGGTCAAGCTGGAGATCGAGTGGGAGCTGCAGAAAGAAAACAGCAACGCTGGAATCCC 4290

Qy 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580

Db 4291 GAAATTCAGTACACTTCCAACACTACAACAAGTCTGTAAATGTGGACTTTTACTGTGGACT 4350

Qy 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599

Db 4351 AATGGCGTGATTTCAGAGCGCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTG 4407

RESULT 10

AF369963

LOCUS

DEFINITION Cloning vector pAAV-RC, 7327 bp DNA circular SYN 22-JUL-2001

ACCESSION AF369963

VERSION AF369963.1 GI:14994089

KEYWORDS Cloning vector pAAV-RC

SOURCE Cloning vector pAAV-RC

ORGANISM artificial sequences; vectors.

REFERENCE 1 (bases 1 to 7327)

AUTHORS Grafsky, A.J. III.

TITLE pAAV-RC for AAV Helper-Free System

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 7327)

AUTHORS Grafsky, A.J. III.

TITLE Direct Submission

JOURNAL Submitted (16-APR-2001) Technical Services, Stratagene, 11011 N. Torrey Pines Rd., La Jolla, CA 92037, USA

FEATURES

Location/Qualifiers

1..7327

/organism="Cloning vector pAAV-RC"

/mol\_type="genomic DNA"

/db\_xref="taxon:161364"

/lab\_host="Escherichia coli"

/note="used for recombinant DNA processing"

131..1996

/note="AAV-2 rep"

CDS

|   |           |   |      |  |  |
|---|-----------|---|------|--|--|
|   |           | <pre>/codon_start=1 /transl_table=11 /product="adeno-associated virus 2 rep protein" /protein_id="AAK76418.1" /db_xref="GI:14994090" /translation="MPGFYEIVIKVPSDLDEHLPGISDSFVNWVAEKEWELPPDSMD LNLEQAPLTVAEKLQRFLEWRVSKAPEALFFVQFEKGESYFHMHLVETTVGVS MVLGRFLSQIREKLIQRIYRGIEPTLPNWFVAVTKRNGAGGNKVVDECYIPNYLLPK TQPELQAWTNMEQYLSACLNLTERKRLVAQHLTHVSQTQEQNKENQNPNSDAPVRS KTSARYMELVGLVDKGITSEKOWIQEDQASYISFNAASNSRSQIKAAALDNAGKIMSL TKTAPDYLVGQQVEDISSNRIYKILELNGYDPQYAAASVFLGWATKKFGKRNITLWLF PATTTGKTNIAEAIHTVFPYGCNVNWTNENFPFNDQVDMVIMWEEGKMTAKVVESAKA ILGSKVRVDQCKSSAQIDPTPIVTSNTNMCAVIDGNSTTFHQQPLQDRMPKFEL TRRLDHDGKVTQKEVKQDFFRWAKDHWVEHEFVYKGGAKKRPAPSDADISEPKRV RESVAQPSSTDAEASINYADRYQNKCSRHVGMNMLMFPCKQCERMNQNSNICFTHGQK DCLECFPVSESQPVSVVKKAYQKLCYIHHIMGKVPDACTACDLNVNLDLDDCIFEQ" 2013..4220 /note="cap VP1" /codon_start=1 /transl_table=11 /product="capsid protein VP1" /protein_id="AAK76419.1" /db_xref="GI:14994091" /translation="MAADGYLPDWELETLSEGIQWMLKPGPPPPKPAERHKDDSRG LVLPGYKLGPFNGLDKGEPVNEADAALEHDKAYDRQLDSDGNPYLYKNHADAEFQE RLKEDTSYGGNLGRAVFOAKRVLPLGLVEEPVKTA PGKRPVEHSPVEPDSSSGTGT KAGQOPARKRLNFGQTGDADSVDPQPLGQPPAAPSPGLTNTMATGSGAPMADNNEGA DGVNSGSNWHCDSTWMDRVITSTRTVALPTYNHLYKQISSQSGASNDNHYFGYS TPWGYPDFNRFCHFSPRDQRLNNWGRPKRLNFKLNIQVKEVTQNDGTTIAN NLSTVQVFTDSEYQLPYVLGSAHQCLPPFPADVFVMPQYGLTLNNGSQAVGRSSF YCLEYFPQMLRTGNFTFSYTFEDVPFHSSYAHSSQSLDRLMNPPLIDQYLYLSRTNY PSGTTTQSRQLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKY HLNGRDSLVPNGPAMASHKDDDEKFFPSGVLIFGKQSGSEKTNVDIEKVMITDEEIR TTNPVATEQYGSVSTNLQRNROAATADVNTQGVLPGMVWQDRDVLQGPWAKIPHPT DGHFHPSPLMGFGFLKHPPQILIKNTVPANPSTTFAAKFASFITQYSTGQVSVEI EWELQKENSXRNWPEIQYTSNYSNVNVDFTVDITNGVYSEPRPIGTRYLTRNL" 4705..5161 /note="fl origin" 5292..6152 /note="confers ampicillin resistance" /codon_start=1 /transl_table=11 /product="beta-lactamase" /protein_id="AAK76420.1" /db_xref="GI:14994092" /translation="MSIQHFRVALIPFFAAFLPVFAHPETLVKVKDAEDQLGARVGY IELDNSGKILESFRPEERFPMSTFKVLLCGAVLSRIDAGEQLGRRIRHYSQNDLVE YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTIGGPKELTAPFLHNMGDHVTRL DRWPELNEAIPNDERDTPMPVAMATLRLKLTGELLTLASRQQLIDWMEADKIVAGPL LRSLAPAGWFIADKSGAGERSGIIAALGPDGKPSRIIVYITGTSQATMDERNRQIA EIGASLIKHW" 6161..7019 /note="Cole1 origin" BASE COUNT 1903 a 1887 c 1912 g 1625 t ORIGIN</pre> |      |  |  |
| Alignment Scores:                             |           |   |      |  |  |
| Pred. No.:                                    | 9.95e-195 | Length:   | 7327 |  |  |
| Score:  | 2764.50   | Matches:  | 496  |  |  |
| Percent Similarity:                           | 89.98%    | Conservative:   | 43   |  |  |
| Best Local Similarity:                        | 82.80%    | Mismatches:   | 59   |  |  |
| Query Match:                                  | 85.04%    | Indels:   | 1    |  |  |
| DB:   | 12        | Gaps:   | 1    |  |  |
| US-09-807-802A-15 (1-599) x AF369963 (1-7327) |           |   |      |  |  |
| Qy  | 1         | ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer  | 20   |  |  |
| Db  | 2424      | ACGGCTCCGGGAAAAAGAGGCCGGTAGAGCACTCTCCTGTGGAGCCAGACTCCTCCTCG   | 2483 |  |  |
| Qy  | 21        | GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly  | 40   |  |  |
| Db  | 2484      | GGAACCGGAAAGCGCGGCAGCAGCCCTGCAAGAAAAAGATTGAATTTTGGTCAGACTGGA  | 2543 |  |  |

|    |      |   |      |
|----|------|---|------|
| Qy | 41   | AspSerGluSerValProAspProGlnProLeuGlyGlyGlyAlaProMetAlaAspAsnAsnGluGly | 60   |
| Db | 2544 | GAGCGAGACTCAGTACCTGACCCCGAGCCTCTCGGACAGCCACCAGCAGCCCTCTGGT            | 2603 |
| Qy | 61   | ValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGly          | 80   |
| Db | 2604 | CTGGGAACATAATACGATGGCTACAGGCAGTGGCGCAATGGCAGACAATAACGAGGCG            | 2663 |
| Qy | 81   | AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp          | 100  |
| Db | 2664 | GCCGACGGAGTGGTAATTCCTCGGAAATTTGGCATTTGGGATTCACATGGATGGGCGAC           | 2723 |
| Qy | 101  | ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr          | 120  |
| Db | 2724 | AGAGTCATCACCCAGCACCCGAAACCTGGGCCCTGCCACCTACAACAACCACCTCTAC            | 2783 |
| Qy | 121  | LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer          | 140  |
| Db | 2784 | AAACAAATTTCCAGCCCAATCA--GGAGCCTCGAACGACAAATCACTACTTTGGCTACAGC         | 2840 |
| Qy | 141  | ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp          | 160  |
| Db | 2841 | ACCCCTTGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCACCACGTGACTGG           | 2900 |
| Qy | 161  | GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe          | 180  |
| Db | 2901 | CAAAGACTCATCAACAACAATGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTT          | 2960 |
| Qy | 181  | AsnIleGlnValLysGluValThrAsnAspGlyValThrThrIleAlaAsnAsnLeu             | 200  |
| Db | 2961 | AACATTCAAGTCAAAGAGGTACGCAGAAATGACGGTACGACGACGATTGCCAATAACCTT          | 3020 |
| Qy | 201  | ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer          | 220  |
| Db | 3021 | ACCAGCAGGTTTCAGGTGTTTACTGACTCGGAGTACCAGTCCCCTACGTTCCCTCGGCTCG         | 3080 |
| Qy | 221  | AlaHisGlnGlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGly             | 240  |
| Db | 3081 | GGCATCAAGGATGCCTCCCGCGTTCCCAGCAGACGTTTCATGGTGGCCACAGTATGGA            | 3140 |
| Qy | 241  | TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu          | 260  |
| Db | 3141 | TACCTCACCCCTGAACAACGGGAGTCAGGCAGTAGGACCGTCTTCTTACTGCTGCGGAG           | 3200 |
| Qy | 261  | TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu          | 280  |
| Db | 3201 | TACTTTCCTTCTCAGATGCTGGCTACCGAAACAACCTTTACCTTCAGCTACACTTTTGAG          | 3260 |
| Qy | 281  | GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro          | 300  |
| Db | 3261 | GACGTTCTCTTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCT          | 3320 |
| Qy | 301  | LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln          | 320  |
| Db | 3321 | CTCATCGACCAGTACCTGTATTACTTGAGCAGACAACAACACTCCAAGTGGAAACCACCA          | 3380 |
| Qy | 321  | AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn          | 340  |
| Db | 3381 | CAGTCAAGGCTTCAGTTTTTCTCAGGCCGGAGCGAGTGACATTCGGGACCAGTCTAGGA           | 3440 |
| Qy | 341  | TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn          | 360  |
| Db | 3441 | TGGCTTCTGGACCCCTGTTACCGCCAGCAGCGAGTATCAAGAAGACATCTGCGGATAACA          | 3500 |
| Qy | 361  | AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle          | 380  |
| Db | 3501 | AACAGTGAATACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTG          | 3560 |
| Qy | 381  | IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet          | 400  |
| Db | 3561 | GTGAATCCGGGCGCGCCATGGCAAGCCCAAGGACGATGAAGAAAAAGTTTTTTTCTCTCAG         | 3620 |

|   |      |   |      |
|---|------|---|------|
| QY  | 401  | SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn  | 420  |
| Db  | 3621 | AGCGGGTTCTCATCTTTGGGAAGCAAGGCTCAGAGAAAAACAAATGTGGACATTGAAAAG  | 3680 |
| QY  | 421  | ValMetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe  | 440  |
| Db  | 3681 | GTCAATGATTACAGACGAAGAGGAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTAT  | 3740 |
| QY  | 441  | GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis  | 460  |
| Db  | 3741 | GGTTCTGTATCTACCAACCTCCAGAGAGGCAACACAGACAGCTACCGCAGATGTCAAC    | 3800 |
| QY  | 461  | AlaMetGlyAlaLeuProGlyMetValTyrGlnAspArgAspValTyrLeuGlnGlyPro  | 480  |
| Db  | 3801 | ACACAAGCGGTTCTTCCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCC   | 3860 |
| QY  | 481  | IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly  | 500  |
| Db  | 3861 | ATCTGGGCAAGATTCCACACACGGACGGACATTTTACCCCTCTCCCTCATGGGTGGA     | 3920 |
| QY  | 501  | PheGlyLeuLysAsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsn  | 520  |
| Db  | 3921 | TTCGGACTTAACACACCTCTCCACAGATTCTCATCAAGAAACACCCCGGTACCTGCGAAT  | 3980 |
| QY  | 521  | ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly  | 540  |
| Db  | 3981 | CCTTCGACCACCTTCAGTGGCGCAAGTTTGCTTCCTTCATCACACAGTACTCCACGGGA   | 4040 |
| QY  | 541  | GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro  | 560  |
| Db  | 4041 | CAGGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAGGAAACAGCAACGCTGGAATCCC     | 4100 |
| QY  | 561  | GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn  | 580  |
| Db  | 4101 | GAAATTCAGTACACTTCCAACTACAACAAGTCTGTTAATGTGGACTTTACTGTGACACT   | 4160 |
| QY  | 581  | AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu     | 599  |
| Db  | 4161 | AATGGCGTGATTTCAGAGCGCTCGCCCCCATTTGGCACCAGATACCTGACTCGTAATCTG  | 4217 |
| RESULT 11                                     |      |   |      |
| AR222045                                      |      |   |      |
| LOCUS   |      |   |      |
| DEFINITION                                    |      |   |      |
| ACCESSION                                     |      |   |      |
| VERSION                                       |      |   |      |
| KEYWORDS                                      |      |   |      |
| SOURCE  |      |   |      |
| ORGANISM                                      |      |   |      |
| REFERENCE                                     |      |   |      |
| AUTHORS                                       |      |   |      |
| TITLE   |      |   |      |
| JOURNAL                                       |      |   |      |
| FEATURES                                      |      |   |      |
| BASE COUNT                                    |      |   |      |
| ORIGIN  |      |   |      |
| Alignment Scores:                             |      |   |      |
| Pred. No.:                                    |      |   |      |
| Score:  |      |   |      |
| Percent Similarity:                           |      |   |      |
| Best Local Similarity:                        |      |   |      |
| Query Match:                                  |      |   |      |
| DB:   |      |   |      |
| US-09-807-802A-15 (1-599) x AR222045 (1-7557) |      |   |      |
| QY  | 1    | ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer  | 20   |
| Db  | 2642 | ACGGCTCCGGGAAAAAAGAGGCCCGGTAGACACTCTCTCTGTGGAGCCAGACTCTCTCTCG | 2701 |

|    |      |  |      |
|----|------|--|------|
| QY | 21   | GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly   | 40   |
| Db | 2702 | GGAACCGGAAAGGGGGCCAGCAGCTTGCAGAAAAGATTGAATTTGGTCAGACTGGA       | 2761 |
| QY | 41   | AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla   | 60   |
| Db | 2762 | GACGACAGCTCAGTACCTGACCCCCCAGCTCTCGGACAGCCACGAGCCCCCTCTGGT      | 2821 |
| QY | 61   | ValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGly      | 80   |
| Db | 2822 | CTGGGAACATAACGATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGGC    | 2881 |
| QY | 81   | AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp   | 100  |
| Db | 2882 | GCCGACGGAGTGGGTAATTCTCTCGGAAATTTGGCATTTCCGATTTCCACATGATGGCGGAC | 2941 |
| QY | 101  | ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr   | 120  |
| Db | 2942 | AGAGTCATCACACACAGCACCCGAAACCTGGGGCCCTGCCACCTACACACACACCTCTAC   | 3001 |
| QY | 121  | LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer   | 140  |
| Db | 3002 | AAACAAATTTCCAGCCAATCA--GGAGCTCGAACGACAATCACTACTTTGGCTACAGC     | 3058 |
| QY | 141  | ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp   | 160  |
| Db | 3059 | ACCCCTTGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTTCACCACGTGACTGG   | 3118 |
| QY | 161  | GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe   | 180  |
| Db | 3119 | CAAGACTCATCAACAACTGTTGGGATTTCCGACCCCAAGAGACTCAACTTCAAGCTCTTT   | 3178 |
| QY | 181  | AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu   | 200  |
| Db | 3179 | AACATTCAAGTCAAAGAGGTCAACGAGATGACGGTACGACGACGATTGCCAATAACCTT    | 3238 |
| QY | 201  | ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer   | 220  |
| Db | 3239 | ACCAGCACGGTTCAGGTGTTTACTGACTCGAGTACCAGTCCCGTACGTCCTCGGCTCG     | 3298 |
| QY | 221  | AlaHisGlnGlyCysLeuProPropheProAlaAspValPheMetIleProGlnTyrGly   | 240  |
| Db | 3299 | GCGCATCAAGGATGCTCCCGCGTTCGCCAGCAGACGTCTTCATGGTCCACAGTATGGA     | 3358 |
| QY | 241  | TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu   | 260  |
| Db | 3359 | TACCTCACCCCTGAACAACGGGAGTCAGGCAGTAGGACGCTCTTCAATTTTACTGCCTGGAG | 3418 |
| QY | 261  | TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu   | 280  |
| Db | 3419 | TACTTTCTCTCAGATGCTGCTACCGGAAACAACCTTTACCTTCAGCTACACTTTTGAG     | 3478 |
| QY | 281  | GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro   | 300  |
| Db | 3479 | GACGTTCTCTTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCT   | 3538 |
| QY | 301  | LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln   | 320  |
| Db | 3539 | CTCATCGACCAGTACCTGTATTACTTGAGCAGAGACAAAACACTCCAAGTGGAAACCACG   | 3598 |
| QY | 321  | AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn   | 340  |
| Db | 3599 | CAGTCAAGGCTTCAGTTTCTCAGGCCGCGGAGTGACATTCGGGACCAGTCTAGGAAC      | 3658 |
| QY | 341  | TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn   | 360  |
| Db | 3659 | TGGCTTCTTGACCCCTGTTACCGCCAGCAGCGAGTATCAAAGACATCTGCGGATACAAAC   | 3718 |
| QY | 361  | AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle   | 380  |
| Db | 3719 | AACAGTGAATACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTG   | 3778 |



QY 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400  
:::|||||  
Db 3779 GTGAATCCGGGCGCCATGGCAAGCCCAAGGACGATGAAGAAAAGTTTTTCTCTCAG 3838  
QY 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420  
|||||:::|||||  
Db 3839 AGCGGGGTTCTCATCTTTGGGAAGCAAGGCTCAGAGAAAACAATGTGGACATTGAAAAG 3898  
QY 421 ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440  
|||||:::|||||  
Db 3899 GTCATGATTACAGACGAAGAGGAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTAT 3958  
QY 441 GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis 460  
|||||:::|||||  
Db 3959 GGTCTGTATCTACCAACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAAC 4018  
QY 461 AlaMetGlyAlaLeuProGlyMetValTyrGlnAspArgAspValTyrLeuGlnGlyPro 480  
Db 4019 ACACAAGGCGTTCTTCCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCC 4078  
QY 481 IleTyrAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500  
Db 4079 ATCTGGGCAAGATTCCACACACGGACGGACATTTTCAACCTCTCCCTCATGGGTGGA 4138  
QY 501 PheGlyLeuLysAsnProProGlnIleLeuLysAsnThrProValProAlaAsn 520  
Db 4139 TTCGGACTTAAACACCCCTCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAAT 4198  
QY 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540  
Db 4199 CCTTCGACCACTTCAGTGGCGGCAAGTTTGCTTCTTCATCACAGTACTCCACGGGA 4258  
QY 541 GlnValSerValGluIleGluTyrGluLeuGlnLysGluAsnSerLysArgTyrAsnPro 560  
Db 4259 CAGGTGAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAAAACAGCAACGCTGGAATCCC 4318  
QY 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580  
Db 4319 GAAATTCAGTACACTTCCAACCTACAACAAGTCTGTAAATGTGGACTTTACTGTGGACACT 4378  
QY 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599  
Db 4379 AATGGCGTGATTACAGAGCCTCGCCCCATTGGCACCAGATACCTTGACTCGTAATCTG 4435

RESULT 12  
AX205073  
LOCUS AX205073 7557 bp DNA linear PAT 30-AUG-2001  
DEFINITION Sequence 3 from Patent WO0155361.  
ACCESSION AX205073  
VERSION AX205073.1 GI:15394314  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1  
AUTHORS Hardy, S.F.  
TITLE Recombinant aav packaging systems  
JOURNAL Patent: WO 0155361-A 3 02-AUG-2001;  
Chiron Corporation (US)  
FEATURES Location/Qualifiers  
source  
1..7557  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"  
/note="recombinant DNA"  
BASE COUNT 1970 a 1940 c 1953 g 1694 t  
ORIGIN  
Alignment Scores:  
pred. No.: 1.04e-194 Length: 7557  
Score: 2764.50 Matches: 496  
Percent Similarity: 89.98% Conservative: 43  
Best Local Similarity: 82.80% Mismatches: 59

Query Match: 85.04% Indels: 1  
DB: 6 Gaps: 1  
US-09-807-802A-15 (1-599) x AX205073 (1-7557)  
QY 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20  
Db 2642 ACGGCTCCGGGAAAAAGAGCGCGGTAGAGCACTCTCTGTGGAGCCAGACTCTCTCTCG 2701  
QY 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40  
Db 2702 GGAACCGAAAGCGCGGCCAGCAGCCTGCAAGAAAAAGATTGAATTTTGGTCAGACTGGA 2761  
QY 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60  
Db 2762 GACGCACTCAGTACCTGACCTCCCGCCAGCCTCTCGACAGCCACCCCTCTCTGGT 2821  
QY 61 ValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnGluGly 80  
Db 2822 CTGGGAACATAATACGATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGC 2881  
QY 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTyrHisCysAspSerThrTyrLeuGlyAsp 100  
Db 2882 GCCGACGAGTGGTAATTCCTCGGGAATTTGGCATTTGCCATTTCCACATGGATGGCGAC 2941  
QY 101 ArgValIleThrThrSerThrArgThrTyrAlaLeuProThrTyrAsnAsnHisLeuTyr 120  
Db 2942 AGAGTCATCACCAACAGCACCGCACTGGGCCCTGGCCCTGACCACTACCAACCACTCTAC 3001  
QY 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140  
Db 3002 AAACAAATTTCCAGCCAATCA--GGAGCCTCGAACGACAATCACTACTTTGGGTACAGC 3058  
QY 141 ThrProTyrGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTyr 160  
Db 3059 ACCCTTGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCCACACGTGACTGG 3118  
QY 161 GlnArgLeuIleAsnAsnAsnTyrGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180  
Db 3119 CAAGACTCATCAACAACAACTGGGATTCGACCCCAAGAGACTCAACTTCAAGCTCTTT 3178  
QY 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrIleAlaAsnAsnLeu 200  
Db 3179 AACATCAAGTCAAAGAGGTCAACGAGATGACGGTACGACGACGATTGCCAATAACCTT 3238  
QY 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220  
Db 3239 ACCAGCAGGTTTCAAGTGTACTGACTCGGAGTACCAGCTCCCGTACGTCTCTCGGCTCG 3298  
QY 221 AlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGly 240  
Db 3299 GCGCATCAAGGATGCCCTCCCGCGTCCAGCAGACGCTTTCATGTGTGCCACAGTATGGA 3358  
QY 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260  
Db 3359 TACCTCACCTGAACAACCGGAGTCAGGCAGTAGGACGCTCTTCAATTTACTGCTGGAG 3418  
QY 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280  
Db 3419 TACTTTCTTCTCAGATGCTGCGTACCGGAAACAACCTTTACCTTTCAGCTACACTTTTGAG 3478  
QY 281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300  
Db 3479 GACGTTCTTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCT 3538  
QY 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320  
Db 3539 CTCATCGACAGTACCTGTATTACTTTGAGCAGACAACAACACTCCCAAGTGGAAACCACCG 3598  
QY 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340  
Db 3599 CAGTCAAGGCTTCAGTTTCTCAGGCGGAGCGAGTGACATTCGGGACCGAGTCTTAGGAAC 3658

```
QY 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrLysThrAspAsnAsn 360
Db 3659 TGGCTTCTCTGGACCCCTGTACCGCCAGCAGCGAGTATCAAGACATCTCGGATAACAAC 3718
QY 361 AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380
Db 3719 AACAGTGAATACTCGTGGAGCTGAGCTACCAAGTACCACTCAATGGCAGAGACTCTCTG 3778
QY 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400
Db 3779 GTGAATCCGGCCCGCCATGGCAAGCCACAGGACGATGAAGAAAGTTTTTCTCTCAG 3838
QY 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420
Db 3839 AGCGGGTTCTCATCTTTGGGAAGCAAGGCTCAGAGAAAAACAATGTGGACATTGAAAG 3898
QY 421 ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440
Db 3899 GTCATGATTACAGACGAAGAGGAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTAT 3958
QY 441 GlyThrValAlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHis 460
Db 3959 GGTTCTGTATCTACCAACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAAC 4018
QY 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 480
Db 4019 ACACAAGGCGTTCTTCCAGGCGATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCC 4078
QY 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500
Db 4079 ATCTGGCAAGATTCCACACACGACGACGACATTTTCACCCCTCTCCCTCATGGGTGGA 4138
QY 501 PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520
Db 4139 TTCGGACTTAAACACCCCTCCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAAT 4198
QY 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540
Db 4199 CCTTCGACCACCTTCAGTCGGCAAGTTTGCTTCCTTCATCACACAGTACTCCACGGGA 4258
QY 541 GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro 560
Db 4259 CAGGTGAGCGTGGAGATCGAGTGGAGCTGCAGAAGGAACAGCAACGCTGGAATCCC 4318
QY 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580
Db 4319 GAAATTCAGTACACTTCCAACTACAACAAGTCTGTTAATGTGGACTTTACTGTGGACACT 4378
QY 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599
Db 4379 AATGGCGTGATTTCAGAGCCTCGCCCAATTGGCACCAGATACCTGACTCGTAATCTG 4435
```

```
RESULT 13
AR222044
LOCUS AR222044 8698 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 2 from patent US 6429001.
ACCESSION AR222044
VERSION AR222044.1 GI:23329400
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 8698)
TITLE Hardy,S.F.
JOURNAL Recombinant AAV packaging systems
FEATURES Patent: US 6429001-A 2 06-AUG-2002;
Location/Qualifiers
source 1. .8698
/organism="unknown"
BASE COUNT 2117 a 2368 c 2306 g 1907 t
ORIGIN
```

Alignment Scores:

```
Pred. No.: 1.25e-194 Length: 8698
Score: 2764.50 Matches: 496
Percent Similarity: 89.98% Conservative: 43
Best Local Similarity: 82.80% Mismatches: 59
Query Match: 85.04% Indels: 1
DB: 6 Gaps: 1
US-09-807-802A-15 (1-599) x AR222044 (1-8698)
QY 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20
Db 2614 ACGGCTCCGGGAAAAAAGAGGGCCGGTAGAGCACTCTCCTGTGGAGCCAGACTCCTCTCG 2673
QY 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40
Db 2674 GGAACCGGAAGGCGGCCAGCAGCCTTCGAAGAAAAAGATTGAATTTTGGTCAGACTGGA 2733
QY 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60
Db 2734 GACGCAGACTCAGTACCTGACCCCCAGCCTCTCGGACAGCCACCAGACGCCCTCTGTT 2793
QY 61 ValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnGluGly 80
Db 2794 CTGGGAACATAACGATGGCTACAGGCACTGGCGCACCAATGGCAGACAATAACGAGGGC 2853
QY 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100
Db 2854 GCCGACGGAGTGGGTAAATTCCTCGGGAATTTGGCATTCGATTCACATGGATGGCGGAC 2913
QY 101 ArgValIleThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120
Db 2914 AGAGTCATCACACCAGCACCCGAACTTGGGCCCTGCCACCTACAACAACACCTCTAC 2973
QY 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140
Db 2974 AAACAAATTTCCAGCCAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGC 3030
QY 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160
Db 3031 ACCCTTGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTTCACCAGCTGACTGG 3090
QY 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180
Db 3091 CAAAGACTCATCAACAACAACACTGGGATTCGACCCCAAGAGACTCAACTTCAAGCTCTTT 3150
QY 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200
Db 3151 AACATTCAAGTCAAAGAGGTACGCAGAAATGACGGTACGACGAGATTGCCAATAACCTT 3210
QY 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220
Db 3211 ACCAGCAGGTTCAGGTGTTTACTGACTCTGGAGTACCAGTCCCCGTACGTCCTCGGCTCG 3270
QY 221 AlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGly 240
Db 3271 GCGCATCAAGGATGCTCCCGCCGTTCCAGCAGACGCTCTTCATGGTGCCACAGTATGGA 3330
QY 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260
Db 3331 TACCTCACCCCTGAACAACGGGAGTCAAGGACGTAGGACGCTCTTCAATTTTACTGCTGGAG 3390
QY 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280
Db 3391 TACTTTCTCTCAGATGCTGCGTACCGGAAACAACACTTACCTTCAGCTACACTTTTGAG 3450
QY 281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300
Db 3451 GACGTTCTCTTCCACAGCAGCTACGCTCACGCCAGAGTCTGGACCGTCTCATGAATCCT 3510
QY 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320
Db 3511 CTCATCGACCAGTACCTGTATTACTTGAGCAGAAACAACACTCCAAAGTGGAAACCACG 3570
```

```
QY 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340
Db 3571 CAGTCAAGCTTCAGTTTCTCAGGCCGAGCGAGTGACATTCGGGACCAGTCTAGGAAC 3630
QY 341 TrpLeuProGlyProCysTyrArgGlnArgValSerLysThrLysThrAspAsnAsn 360
Db 3631 TGGCTTCTGGACCTGTTACCGCCAGCAGCGAGTATCAAGACATCTCGGATACAAAC 3690
QY 361 AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380
Db 3691 AACAGTGAATACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTG 3750
QY 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePhePromet 400
Db 3751 GTGAATCCGGGCCCGCCCATGGCAAGCCACAGGACGATGAAGAAAAGTTTTTTCCTCAG 3810
QY 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420
Db 3811 AGCGGGGTTCTCATCTTTGGGAAGCAAGGCTCAGAGAAACAAATGTGGACATTTGAAAAG 3870
QY 421 ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440
Db 3871 GTCATGATTACAGACGAAGAGAAATCAGACAACCAATCCCGTGGCTACGGAGCAGTAT 3930
QY 441 GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis 460
Db 3931 GGTTCGTATCTACCAACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAAC 3990
QY 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTrpLeuGlnGlyPro 480
Db 3991 ACACAAGGCGTTCTCCAGGCATGGTCTGGCAGGACAGAGATGTACCTTCAGGGGCC 4050
QY 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500
Db 4051 ATCTGGGCAAGATTCACACACGACGGACGACATTTTCACCCCTCTCCCTCATGGGTGA 4110
QY 501 PheGlyLeuLysAsnProProGlnIleLeuLysAsnThrProValProAlaAsn 520
Db 4111 TTCGGACTTAACACACCTCCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAAT 4170
QY 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540
Db 4171 CCTTCGACCACCTTCAGTGGCGCAAGTTTGCTTCTTCATCACACAGTACTCCACGGGA 4230
QY 541 GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro 560
Db 4231 CAGGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAAACAGCAACGCTGGAATCCC 4290
QY 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580
Db 4291 GAAATTCAGTACACTTCCAACACTACAAGTCTGTTAATGTGGACITTTACTGTGGACACT 4350
QY 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599
Db 4351 AATGGCGTGATTACAGAGCCTCGCCCCATTGGCACCAGATACCTGACTCGTAATCTG 4407
```

```
RESULT 14
AX205072 LOCUS 8698 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 2 from Patent WO0155361.
ACCESSION AX205072
VERSION AX205072.1 GI:15394313
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Hardy,S.F.
TITLE Recombinant aav packaging systems
JOURNAL Patent: WO 0155361-A 2 02-AUG-2001;
Chiron Corporation (US)
FEATURES
source 1. .8698
```

```
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
/note="recombinant DNA"

BASE COUNT 2117 a 2368 c 2306 g 1907 t
ORIGIN

Alignment Scores:
Pred. No.: 1.25e-194 Length: 8698
Score: 2764.50 Matches: 496
Percent Similarity: 89.98% Conservative: 43
Best Local Similarity: 82.80% Mismatches: 59
Query Match: 85.04% Indels: 1
DB: 6 Gaps: 1

US-09-807-802A-15 (1-599) x AX205072 (1-8698)

QY 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20
Db 2614 ACGGCTCCGGGAAAAAAGAGCGCGGTAGAGCACTCTCTGTGGAGCCAGACTCCTCTCG 2673
QY 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40
Db 2674 GGAACCGGAAGGGCGGCCAGCAGCCTGCAAGAAAAAGATTGAATTTGGTCAGACTGGA 2733
QY 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60
Db 2734 GACGAGACTCAGTACTGACCCCGCAGCCTCTCGGACAGCCACCAGCAGCCCCCTCTGGT 2793
QY 61 ValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGly 80
Db 2794 CTGGGAACATAATACGATGGCTACAGGCAGTGCGCGCAACCAATGGCAGACAATAACGAGGC 2853
QY 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100
Db 2854 GCCGACGGAGTGGGTAAATCTCGGAAATTTGGCAATTCGCAATTCACATGGATGGCGAC 2913
QY 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120
Db 2914 AGAGTCATCACCAACAGCAGCAGCCGAAACCTGGGCCCTGCCACCTTACCAACACCACCTCTAC 2973
QY 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140
Db 2974 AAACAAATTTCCAGCCCAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGC 3030
QY 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160
Db 3031 ACCCTTGGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTACCACCGTACTGG 3090
QY 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180
Db 3091 CAAAGACTCATCAACAACAACTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTT 3150
QY 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrIleAlaAsnAsnLeu 200
Db 3151 AACATTCAAGTCAAGAGGTCACGCAGAAATGACGGTACGCGTACGACGATTCCAATAACCTT 3210
QY 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220
Db 3211 ACCAGCAGCGTTTCAGGTGTTTACTGACTCGGAGTACCAGTCCCGTACGTCTCTCGGCTCG 3270
QY 221 AlaHisGlnGlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGly 240
Db 3271 GCGCATCAAGGATGCCCTCCCGCGTTCCAGCAGACGCTCTTCATGGTGCCACAGTATGGA 3330
QY 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260
Db 3331 TACCTCACCTGAACAACGGGAGTCAGGCAGTAGGACGCTCTTTCATTTACTGCTGGAG 3390
QY 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280
Db 3391 TACTTTCCTTCTCAGATGCTGCGTACCGGAAACAACCTTTTACCTTACGCTACACTTTTGAG 3450
```



|    |      |   |      |
|----|------|---|------|
| QY | 281  | GluValProPheHisSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro     | 300  |
| Db | 3451 | GACGTTCCCTTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCT  | 3510 |
| QY | 301  | LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln  | 320  |
| Db | 3511 | CTCATCGACCACTGATGATTACTTGAGCAGACAAACACTCCAAGTGAACCAACCACG     | 3570 |
| QY | 321  | AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn  | 340  |
| Db | 3571 | CAGTCAAGGCTTCAGTTTCTCAGGCCGGAGCGAGTGACATTCCGGAGCACAGTCTAGGAAC | 3630 |
| QY | 341  | TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn  | 360  |
| Db | 3631 | TGGCTTCCTGGACCCCTGTTACCGCCAGCAGCGAGTATCAAAGACATCTCCGGATAACAAC | 3690 |
| QY | 361  | AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle  | 380  |
| Db | 3691 | AACAGTGAATACTCGTGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTG   | 3750 |
| QY | 381  | IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet  | 400  |
| Db | 3751 | GTGAATCCGGCCCGCCGTCATGGCAAGCCACCAAGGAGCATGAAGAAAGTTTTCCTCAG   | 3810 |
| QY | 401  | SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn  | 420  |
| Db | 3811 | AGCGGGTTCTCATCTTGGGAAGCAAGGCTCAGAGAAACAAATGTGGACATTGAAAG      | 3870 |
| QY | 421  | ValMetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe  | 440  |
| Db | 3871 | GTCATGATTACAGACGAAGAGGAAATCAGGACAACCAATCCCGTGGTACCGAGCAGTAT   | 3930 |
| QY | 441  | GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis  | 460  |
| Db | 3931 | GGTTCGTGATCTACCAACTCCAGAGAGGCAACAGACAGCAGCTACCGCAGATGTCAAC    | 3990 |
| QY | 461  | AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro  | 480  |
| Db | 3991 | ACACAAGGCGTTCTTCCAGGCATGGTCTGGCAGGACAGAGATGTACCTTCAGGGGCC     | 4050 |
| QY | 481  | IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly  | 500  |
| Db | 4051 | ATCTGGCAAGATTCCACACACGACGACGACATTTTCACCCCTCTCCCTCATGGGTGA     | 4110 |
| QY | 501  | PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn     | 520  |
| Db | 4111 | TTCGGACTTAACACCCCTCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAAT   | 4170 |
| QY | 521  | ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly  | 540  |
| Db | 4171 | CCTTCGACCACTTCAGTGGGCAAGTTTGCTTCCTTCATCACACAGTACTCCACGGGA     | 4230 |
| QY | 541  | GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro  | 560  |
| Db | 4231 | CAGGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAAACACAGAAACGCTGGAATCCC  | 4290 |
| QY | 561  | GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn  | 580  |
| Db | 4291 | GAAATTCAGTACACTTCCAACATCAACAAGTCTGTTAATGTGGACTTTTACTGTGGACACT | 4350 |
| QY | 581  | AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu     | 599  |
| Db | 4351 | AATGGCGTGATTTCAGAGCCTCGCCCCATTGGCACCCAGATACCTGACTCGTAATCTG    | 4407 |

RESULT 15

AY243021  
LOCUS  
DEFINITION Non-human primate Adeno-associated virus isolate AAVch.5 capsid protein (VP1) gene, complete cds.  
ACCESSION  
VERSION AY243021.1 GI:29650537  
KEYWORDS  
SOURCE Non-human primate Adeno-associated virus

|   |  |  |             |
|---|--|--|-------------|
| ORGANISM                                      | Non-human primate Adeno-associated virus   |  |             |
| REFERENCE                                     | Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.  |  |             |
| AUTHORS                                       | 1 (bases 1 to 2208)<br>Gao,G., Alvira,M.R., Somanathan,S., Lu,Y., Vandenbergh,L.H.,<br>Rux,J.J., Calcedo,R., Sanmiguel,J., Abbas,Z. and Wilson,J.M.  |  |             |
| TITLE   | Adeno-associated viruses undergo substantial evolution in primates during natural infections   |  |             |
| JOURNAL                                       | Proc. Natl. Acad. Sci. U.S.A. 100 (10), 6081-6086 (2003)   |  |             |
| PUBMED  | 12716974   |  |             |
| REFERENCE                                     | 2 (bases 1 to 2208)<br>Gao,G., Alvira,M.R., Somanathan,S., Lu,Y., Vandenbergh,L.H.,<br>Rux,J., Calcedo,R., Sanmiguel,J., Abbas,Z. and Wilson,J.M.  |  |             |
| AUTHORS                                       | Direct Submission  |  |             |
| TITLE   | Submitted (23-FEB-2003) Department of Medicine, University of  |  |             |
| JOURNAL                                       | Pennsylvania School of Medicine and Wistar Institute, 204 Wistar   |  |             |
|   | Institute, 3601 Spruce Street, Philadelphia, PA 19104, USA   |  |             |
| FEATURES                                      | Location/Qualifiers  |  |             |
| source  | 1..2208  |  |             |
|   | /organism="Non-human primate Adeno-associated virus"   |  |             |
|   | /proviral  |  |             |
|   | /mol type="genomic DNA"  |  |             |
|   | /isolate="AAVch.5"   |  |             |
|   | /db_xref="taxon:226582"  |  |             |
| gene  | 1..2208  |  |             |
|   | /gene="VP1"  |  |             |
| CDS   | 1..2208  |  |             |
|   | /gene="VP1"  |  |             |
|   | /codon_start=1   |  |             |
|   | /product="capsid protein"  |  |             |
|   | /protein_id="AAO88207.1"   |  |             |
|   | /db_xref="GI:29650538"   |  |             |
|   | /translation="MAADGYLPDWLEDTLSEGIQWVKLKPGRPPPKPNQHRDSDRG<br>LVLPGYKYLGFNGLDKGEFVNEADAALHDKAYDHLKQGDNPYLKYNHADAEPQE<br>RLQEDTSFGNLRGAVFOAKRVLEPLGLVEEAVKAPGKKRPIEQSPAEPDSSGGIG<br>KSGQQPAKKRLNFGQTGDTESVDPQPIGEPPAAPSGVSNMTMASGGGAPMADNNEGA<br>DGVNSSGNWHCDSTWMDRVITTSRTWALPTYNNHLYKQISSSEGATNDNHYFGYS<br>TPWGYDFNRFHCFSPRDWQRLINNNWFRPKLNFKLNIQVKEVTQNDGTTTIAN<br>NLTSTVQVFTDSEYQLPYVLGSAHQGLPPFPADVMIPOYGYLTJLNNGSQAVGRSSF<br>YCLEYFPSQLRTGNFTFSYTFEDVPFHSSYAHQSLLDRLMNPILIDQLYLYLSKTQG<br>TSGTTQSRQLQFSQAGPSSMAQAKNWLPGPSYRQQRMSKTANDNNNSEFAWTAATKY<br>YLNGRNSLVNPGPPPMASHKDEEKYFPMHGNLIFGKQGTGTNNVDIESVLITDEBEIR<br>TTNPVATEQYQGVATNHQSQNTTASGVSQGLPGMWQDRDVLQGPWAKTPHT<br>DGHFHPSPLMGGFGLKHPPOILIKNTPVPANPATFTPGKFAFITQYSTGQVSVEI<br>EWELQKENSKRWNPEIQYTSYNKSVNVEFTVDANGVYSEPRPIGTRYLTRNL" |  |             |
| BASE COUNT                                    | 617 a  | 638 c  | 518 g 435 t |
| ORIGIN  |  |  |             |
| Alignment Scores:                             |  |  |             |
| Pred. No.:                                    | 3.3e-195-  | Length:  | 2208        |
| Score:  | 2761.50  | Matches:   | 497         |
| Percent Similarity:                           | 89.98%   | Conservative:  | 42          |
| Best Local Similarity:                        | 82.97%   | Mismatches:  | 59          |
| Query Match:                                  | 84.94%   | Indels:  | 1           |
| DB:   | 14   | Gaps:  | 1           |
| US-09-807-802A-15 (1-599) x AY243021 (1-2208) |  |  |             |
| QY  | 1  | ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSer    | 20          |
| Db  | 412  | ACGGCTCCTGGAAAAAGAGACCTATAGACAGTCTCCTGCAGAACCGGACTCTTCTCG    | 471         |
| QY  | 21   | GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly | 40          |
| Db  | 472  | GGCATCGGCAATCAGGCCAGCAGCCCGCTAAGAAAAGACTCAATTTTGGTCAGACTGGC  | 531         |
| QY  | 41   | AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla | 60          |
| Db  | 532  | GACACAGAGTCAGTCCAGACCCCTCAACCAATCGGAGAACCCCGCAGCCCTCTGTT     | 591         |
| QY  | 61   | ValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnGluGly       | 80          |
| Db  | 592  | GTGGGATCTAATACAAATGGCTTCAGGCGGTGGGCGACCAATGGCAGACAATAACGAAGC | 651         |

Qy 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100  
Db 652 GCCACGGAGTGGGTAATTCCTCGGAAATGGCATTCGATTCCACATGGATGGGCGAC 711  
Qy 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120  
Db 712 AGAGTTATCACCACCAGCACACAAGAACCTGGGCCCTCCCCACCTACAATAATCACCTCTAC 771  
Qy 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140  
Db 772 AAGCAAATCTCCAGCGAATCG---GGAGCCACCAACGACCAACCACTACTTCGGCTACAGC 828  
Qy 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160  
Db 829 ACCCCCTGGGGGTATTTTGACTTTTAAACAGATTCCACTGTCTCTCACCACCGTACTGG 888  
Qy 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180  
Db 889 CAGCGACTCATCAACAACAACTGGGGATTTAGACCCCAAGAACTCAATTTCAAGCTCTTC 948  
Qy 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200  
Db 949 AACATCCAAAGTCAAGGAGGTCACGCGAATGATGGAACCAACGACCATCGCCAATAACCTT 1008  
Qy 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220  
Db 1009 ACCAGCACGGTGCAGGTCTTCACAGACTCTGAGTACCAGCTGCCTACGTCTCTCGGTTG 1068  
Qy 221 AlaHisGlnGlyCysLeuProPropheProAlaAspValPheMetIleProGlnTyrGly 240  
Db 1069 GCTCACCAAGGCTGCCTTCCGCCGTTCCAGCAGACGTCTTCATGATTCTCTCAGTACGGC 1128  
Qy 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260  
Db 1129 TACTTGACTCTGAACAATGGCAGCCCAAGCGGTAGGACGTTCTTCATTCTACTGTCTAGAG 1188  
Qy 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280  
Db 1189 TATTTTCCCTCTCAGATGCTGAGGACGGGAAACAACCTTCACCTTCAGCTACACTTTTGAA 1248  
Qy 281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300  
Db 1249 GACGTGCCTTTCCACAGCAGCTACGCGCAAGCCAGAGTCTGGATCGGCTGATGAATCCT 1308  
Qy 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320  
Db 1309 CTCATTGACCAGTACCTGTATTACCTGAGCAAAACTCAGGGTACAAGTGAACAACCGCAG 1368  
Qy 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340  
Db 1369 CAATCGAGACTGCAGTTTCAGCCAAAGCTGGGCCTAGCTCCATGGCTCAGCAGGCCAAAAC 1428  
Qy 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360  
Db 1429 TGGCTACCGGGACCCAGCTACCGACAGCAGCGAATGTCTAAGACGGCTAATGACAACAAC 1488  
Qy 361 AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380  
Db 1489 AACAGTGAATTTGGCTTGGACTGCAGCCACCAAAATATTACCTGAATGGAAGAAATTTCTCTG 1548  
Qy 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400  
Db 1549 GTCAATCCCCGGGCCCCCAATGGCCAGTCAACAAGGACGATGAGGAAAAGTATTTCCCCATG 1608  
Qy 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420  
Db 1609 CACGGAAATCTCATCTTTGGAAACAAGGCCACAGGAACCTACCAATGTGGACATTGAATCA 1668  
Qy 421 ValMetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440  
Db 1669 GTGCTTATTACAGACGAAGAAGAAATCAGAACAACTAATCTCTGTGGCTACAGAACAATAC 1728  
Qy 441 GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis 460

Db 1729 GGACAGGTTGCCAACCAACCATCAGAGTCAGAAACACACAGCTTCCTATGGAAGTGTGGAC 1788  
Qy 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 480  
Db 1789 AGCCAGGGAATCTTACCTGGAATGGTGTGGACGACCGCATGTCTATCTTCAAGGTCCC 1848  
Qy 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500  
Db 1849 ATTTGGGCCAAAACCTCTCACACGACGGACACTTTCATCTCTCCGCTCATGGAGGC 1908  
Qy 501 PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520  
Db 1909 TTTGGACTGAACACACCCTCTCCCCAGATCCTGATCAAAAAACACACCTGTGCCAGCAAT 1968  
Qy 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540  
Db 1969 CCGCGACCACTTTCACCTCTCGAAAGTTTGTCTTCGTTTCATTTACCCAGTATTCACCGGA 2028  
Qy 541 GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro 560  
Db 2029 CAGGTCAGCGTGGAAATAGAGTGGGAGCTGCAGAAAGAAACAGCAACCGCTGGAAACCCA 2088  
Qy 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580  
Db 2089 GAAATTCAGTACACCTCCAACCTACAACAAGTCGGTGAATGTGGAGTTTACCCGTGGACGCA 2148  
Qy 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599  
Db 2149 AACGGTGTATTCTGAACCCCGCCCTATTGGCACTCGTTACCTTACCCGGAACCTTG 2205

Search completed: February 14, 2004, 22:40:21  
Job time : 6271.04 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 14, 2004, 15:18:00 ; Search time 410.871 Seconds  
(without alignments)  
3935.452 Million cell updates/sec

Title: US-09-807-802A-15  
Perfect score: 3251  
Sequence: 1 TAPGKKRPVEQSPQEPDSSS.....NNGLYTEPRPIGTRYLTRPL 599

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09807802/runat\_11022004\_175606\_15919/app\_query.fasta\_1.2389  
-DB=N Geneseq\_19Jun03 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09807802@CGN\_1\_1\_806@runat\_11022004\_175606\_15919 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq\_19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*
- 25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match % | Length | DB | ID        | Description        |
|------------|--------|---------------|--------|----|-----------|--------------------|
| 1          | 3251   | 100.0         | 1800   | 21 | AAD000778 | Adeno-associated v |
| 2          | 3251   | 100.0         | 2211   | 21 | AAD000777 | Adeno-associated v |
| 3          | 3251   | 100.0         | 4718   | 21 | AAD000772 | Adeno-associated v |
| 4          | 3229   | 99.3          | 4683   | 22 | AAF23749  | AAV6 DNA sequence. |
| 5          | 2906   | 89.4          | 1605   | 21 | AAD000779 | Adeno-associated v |
| 6          | 2832   | 87.1          | 4722   | 22 | AAF23748  | AAV3B DNA sequence |
| 7          | 2801   | 86.2          | 8178   | 21 | AAD00834  | Chimeric adeno-ass |
| 8          | 2764.5 | 85.0          | 4072   | 22 | AAH26327  | Adenovirus helper  |
| 9          | 2764.5 | 85.0          | 4679   | 22 | AAI66974  | Adeno-associated v |
| 10         | 2764.5 | 85.0          | 4679   | 22 | AAF23750  | AAV2 DNA sequence. |
| 11         | 2764.5 | 85.0          | 4679   | 24 | ABK89694  | Adeno-associated v |
| 12         | 2764.5 | 85.0          | 4679   | 25 | ABV76133  | Adeno-associated v |
| 13         | 2764.5 | 85.0          | 7557   | 22 | AAH26326  | Adeno-associated v |
| 14         | 2764.5 | 85.0          | 8698   | 22 | AAH26325  | Wild-type adeno-as |
| 15         | 2722.5 | 83.7          | 4680   | 17 | AAT09008  | Wild-type adeno-as |
| 16         | 2722.5 | 83.7          | 4680   | 25 | ABX14497  | Wild-type adeno-as |
| 17         | 2706.5 | 83.3          | 4675   | 22 | AAH41481  | Adeno-associated v |
| 18         | 2706.5 | 83.3          | 4675   | 22 | AAF89931  | Nucleotide sequenc |
| 19         | 2706.5 | 83.3          | 4675   | 24 | ABS69879  | Human adeno-associ |
| 20         | 2706.5 | 83.3          | 4675   | 24 | ABS69880  | Human adeno-associ |
| 21         | 2706.5 | 83.3          | 4675   | 24 | ABA02989  | Adeno-associated v |
| 22         | 2398.5 | 73.8          | 8151   | 21 | AAD00832  | Chimeric adeno-ass |
| 23         | 1830.5 | 56.3          | 1800   | 19 | AAV21657  | AAV4 VP2 coat prot |
| 24         | 1830.5 | 56.3          | 1800   | 25 | ABX15816  | Adeno-associated v |
| 25         | 1830.5 | 56.3          | 2208   | 19 | AAV21650  | Adeno-associated v |
| 26         | 1830.5 | 56.3          | 2208   | 25 | ABX15809  | AAV4 VP1 capsid pr |
| 27         | 1830.5 | 56.3          | 4767   | 19 | AAV21648  | Adeno-associated v |
| 28         | 1830.5 | 56.3          | 4767   | 25 | ABX15807  | AAV4 genome. Aden  |
| 29         | 1820.5 | 56.0          | 7215   | 21 | AAD00831  | Adeno-associated v |
| 30         | 1719   | 52.9          | 2487   | 18 | AAT69854  | Hybrid adeno-assoc |
| 31         | 1701   | 52.3          | 4570   | 16 | AAT01190  | Duck parvovirus ca |
| 32         | 1700.5 | 52.3          | 2264   | 21 | AAZ49213  | Barbary duck parvo |
| 33         | 1700.5 | 52.3          | 2264   | 21 | AAZ49214  | Adeno associated v |
| 34         | 1700.5 | 52.3          | 2307   | 21 | AAZ49212  | Adeno associated v |
| 35         | 1700.5 | 52.3          | 4652   | 21 | AAZ49209  | Adeno associated v |
| 36         | 1700.5 | 52.3          | 4652   | 24 | AAS17712  | Adebo-associated v |
| 37         | 1690.5 | 52.0          | 1617   | 19 | AAV21658  | AAV4 VP3 coat prot |
| 38         | 1690.5 | 52.0          | 1617   | 25 | ABX15817  | Adeno-associated v |
| 39         | 772    | 23.7          | 2271   | 21 | AAD00833  | Adeno-associated v |
| 40         | 497.5  | 15.3          | 2343   | 20 | AAH81583  | Erythrovirus V9 DN |
| 41         | 497.5  | 15.3          | 4677   | 16 | AAT49535  | Human parvovirus 9 |
| 42         | 497.5  | 15.3          | 5028   | 20 | AAH81580  | Genomic DNA sequen |
| 43         | 493    | 15.2          | 2380   | 25 | ABZ59573  | Human parvovirus B |
| 44         | 493    | 15.2          | 2380   | 25 | ABZ59576  | Human parvovirus B |
| 45         | 493    | 15.2          | 4678   | 25 | ABZ59570  | Human parvovirus B |

ALIGNMENTS

RESULT 1  
AAD00778  
ID AAD00778 standard; DNA; 1800 BP.  
XX  
AC AAD00778;  
XX  
DT 08-SEP-2000 (first entry)  
XX  
DE Adeno-associated virus serotype 1 capsid protein VP2 DNA.  
XX  
KW Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;  
KW cap protein; recombinant viral vector; gene delivery; gene therapy;  
KW vaccine; transgene; VP2; ss.  
XX  
OS Adeno associated virus serotype 1.  
XX  
FH Key Location/Qualifiers



```
FT CDS 1..1800
FT /*tag= a
FT /product= "VP2 protein"
FT /partial
XX
PN WO200028061-A2.
XX
PD 18-MAY-2000.
XX
PF 02-NOV-1999; 99WO-US25694.
XX
PR 05-NOV-1998; 98US-0107114.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Wilson JM, Xiao W;
XX
DR WPI; 2000-376571/32.
XX
DR P-PSDB; AAY71168.
XX
PT Novel adeno-associated virus serotype 1 polynucleotide useful for
PT preparation of medicament for delivery of a transgene to a host
XX
PS Claim 10; Page 90-93; 108pp; English.
XX
CC The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA
CC which is characterised by two inverted terminal repeats (ITR) and open
CC reading frames for rep and capsid (cap) proteins. The rep reading frame
CC encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap
CC reading frame encodes three structural proteins, VP1, VP2 and VP3.
CC The AAV-1 sequence or its fragments particularly ITRs, rep and cap
CC coding regions, are useful in production of recombinant viral vectors
CC for gene delivery. These vectors can be used as gene therapy
CC vectors, vaccine vectors or antisense delivery vectors. The AAV-1 does
CC not induce the formation of neutralising antibodies specific to any
CC serotype of AAV hence is useful for transforming host cells, and in the
CC preparation of a medicament for the delivery of transgene to a host.
CC The present sequence is an AAV-1 DNA encoding a cap protein VP2 which is
CC useful in the production of recombinant viral vector for gene delivery.
XX
SQ Sequence 1800 BP; 483 A; 532 C; 409 G; 376 T; 0 other;

Alignment Scores:
Pred. No.: 1.09e-259 Length: 1800
Score: 3251.00 Matches: 599
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

US-09-807-802A-15 (1-599) x AAD00778 (1-1800)
QY 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20
Db 1 ACGGCTCCTGGAAAGAAACGTCCGGTAGAGCAGTCGCCACAAAGAGCCAGACTCCTCCTCG 60
QY 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40
Db 61 GGCATCGGCAAGACAGGCCAGCAGCCCGCTAAAGAGAGACTCAATTTTGGTCAGACTGGC 120
QY 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60
Db 121 GACTCAGAGTCAGTCCCGGATCCACAACCTCTCGAGAACCTCCAGCAACCCCGCTGCT 180
QY 61 ValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGly 80
Db 181 GTGGGACCTACTACAAATGGCTTCAGGCGGTGGCCACCAATGGCAGACATAACGAAGGC 240
QY 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100
Db 241 GCCGACGGAGTGGGTAAATGCCTCAGGAAATTGGCAATTGCGATTCCACATGGCTGGCGAC 300
QY 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTrpAsnAsnHisLeuTyr 120
```

```
Db 301 AGAGTCATCACCACGACCCGACCTGGGCGCTTGGCCACCTACAATAACCACTCTAC 360
QY 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140
Db 361 AAGCAAAATCTCCAGTGCTTCAACGGGGGCCAGCAACGACCACTACTTCGGGTACAGC 420
QY 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160
Db 421 ACCCCCTGGGGTATTTTGATTTCAACAGATTCCACTGCCACTTTTTCACCACGTGACTGG 480
QY 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180
Db 481 CAGCGACTCATCAACAACAATTTGGGATTTCCGGGCCCAAGAGACTCAACTTCAAACTCTTC 540
QY 181 AsnIleGlnVallysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200
Db 541 AACATCCAAGTCAAGGAGGTCAAGCAATGATGGCGTCACAACCATCGCTAATAACCTT 600
QY 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220
Db 601 ACCAGCAGGTTCAAGTCTTCTCGGACTCGGAGTACCAGCTTCCGTACGTCCTCGGCTCT 660
QY 221 AlaHisGlnGlyCysLeuProPropheProAlaAspValPheMetIleProGlnTyrGly 240
Db 661 GCGCACCCAGGGCTGCCCTCCCTCCGTTCCCGCGGACGTGTTCAATGATTCGCAATACGGC 720
QY 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260
Db 721 TACTGACGCTCAACAATGGCAGCCCAAGCCGTGGAGCTTCATCTCTTTTACTGCTGGAA 780
QY 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280
Db 781 TATTTCCCTTCTCAGATGCTGAGAACGGGCAACAACCTTTACCTTCAGCTACACCTTTGAG 840
QY 281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300
Db 841 GAAGTGCCTTTCACAGCAGCTACGCGCACAGCCAGAGCTGGACCGCTGATGAATCCT 900
QY 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320
Db 901 CTCATCGACCAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTCCCAA 960
QY 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340
Db 961 AACAAAGGACTTGCTGTTTAGCGGTGGGTCTCCAGCTGGCATGTCTGTTACGCCCAAAAC 1020
QY 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360
Db 1021 TGGCTACCTGGACCCCTGTTATCGGCAGCAGCGCGTTCCTAAAACAAACAGACAAAC 1080
QY 361 AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380
Db 1081 AACAGCAATTTTACCTGGACTGGTGTCTTCAAAATATATACCTCAATGGCGTGAATCCATC 1140
QY 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400
Db 1141 ATCAACCTTGGCACTGCTATGGCTCACACAAAGACGACGAGACAAAGTTCCTTTCCCATG 1200
QY 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420
Db 1201 AGCGGTGTCATGATTTTGGAAAAGAGAGCGCCGGAGCTTCAAAACACTGCATTTGGACAAT 1260
QY 421 ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440
Db 1261 GTCATGATTACAGACGAAGAGGAAATTAAGCCACTAACCTGTGGCCACCGAAAGATTT 1320
QY 441 GlyThrValAlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHis 460
Db 1321 GGGACCGTGGCAGTCAATTTCCAGAGCAGCAGCAGACCTGCGACCGGAGATGTGCAT 1380
QY 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 480
```

Db 1381 GCTATGGAGCATTACCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCAGGGTCCC 1440

Qy 481 ILeTrrAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500

Db 1441 ATTTGGGCCAAATTCCTCACACAGATGGACACTTTTCACCCGCTCTCCTCTTATGGCGGC 1500

Qy 501 PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520

Db 1501 TTTGGACTCAAGAACCCGCCCTCCTCAGATCCTCATCAAAACACAGCCTGTTCTCTGCGAAT 1560

Qy 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540

Db 1561 CCTCCGGCGAGTTTTCAGCTACAAAGTTTGCTTCATTCATCACCCCAATACTCCACAGGA 1620

Qy 541 GlnValSerValGluIleGluTrrPgluLeuGlnLysGluAsnSerLysArgTrrPasnPro 560

Db 1621 CAAGTGAGTGTGGAAATTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCC 1680

Qy 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580

Db 1681 GAAGTGCAGTACACATCCAAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAC 1740

Qy 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599

Db 1741 AATGGACTTATACTGAGCCTCGCCCATTTGGCACCCCGTTACCTTACCCGTCCCTG 1797

RESULT 2  
AAD00777

ID AAD00777 standard; DNA; 2211 BP.

XX AAD00777;

AC AAD00777;

XX 08-SEP-2000 (first entry)

DE Adeno-associated virus serotype 1 capsid protein VP1 DNA.

XX Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;

KW cap protein; recombinant viral vector; gene delivery; gene therapy;

KW vaccine; transgene; VP1; ss.

XX Adeno associated virus serotype 1.

OS Adeno associated virus serotype 1.

XX Key Location/Qualifiers

FT CDS 1..2211

FT /\*tag= a

FT /product= "VP1 protein"

XX WO200028061-A2.

XX 18-MAY-2000.

XX 02-NOV-1999; 99WO-US25694.

XX 05-NOV-1998; 98US-0107114.

XX (UYPE-) UNIV PENNSYLVANIA.

PI Wilson JM, Xiao W;

XX WPI; 2000-376571/32.

DR P-PSDB; AAY71167.

XX Novel adeno-associated virus serotype 1 polynucleotide useful for

PT preparation of medicament for delivery of a transgene to a host

XX Claim 10; Page 83-87; 108pp; English.

XX The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA

CC coding regions, are useful in production of recombinant viral vectors

CC for gene delivery. These vectors can be used as gene therapy

CC vectors, vaccine vectors or antisense delivery vectors. The AAV-1 does

CC not induce the formation of neutralising antibodies specific to any

CC serotype of AAV hence is useful for transforming host cells, and in the

CC preparation of a medicament for the delivery of transgene to a host.

CC The present sequence is an AAV-1 DNA encoding a cap protein VP1 which is

CC useful in the production of recombinant viral vector for gene delivery.

XX SQ Sequence 2211 BP; 577 A; 652 C; 538 G; 444 T; 0 other;

Alignment Scores:

Pred. No.: 1.44e-259 Length: 2211

Score: 3251.00 Matches: 599

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 21 Gaps: 0

US-09-807-802A-15 (1-599) x AAD00777 (1-2211)

Qy 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20

Db 412 ACGGCTCCTGGAAAGAAACGTCCGGTAGAGAGTCGCCACAAGAGCCAGACTCTCTCTCG 471

Qy 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40

Db 472 GGCATCGGCAAGACAGGCCAGCCCCCGCTAAAGAGAGACTCAATTTTGGTCAGACTGGC 531

Qy 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60

Db 532 GACTCAGAGTCAGTCCCGCATCCACAACCTCTCGGAGAACCTCCAGCAACCCCGCTGCT 591

Qy 61 ValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGly 80

Db 592 GTGGGACCTACTACATGGCTTCAGGGCGTGGCGCACCATTGGCAGACAATAACGAAGGC 651

Qy 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrrPgluLeuGlyAsp 100

Db 652 GCCGACGGAGTGGGTAATGCCTCAGGAAATTGGCATTCGATTCACATGGCTGGCGGAC 711

Qy 101 ArgValIleThrThrSerThrArgThrTrrPgluLeuProThrTyrAsnAsnHisLeuTyr 120

Db 712 AGAGTCATCACCCAGCACCCGACCTGGGCCCTTGGCCACCTACATAACCCACTCTAC 771

Qy 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140

Db 772 AAGCAAATCTCCAGTGTCTTCAACGGGGCCAGCAACGACAACCACTACTTTCGGCTACAGC 831

Qy 141 ThrProTrrPgluTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrr 160

Db 832 ACCCCCTGGGGTATTTTGATTTCAACAGATTCCACTGCCACTTTTTCACCACGTGACTGG 891

Qy 161 GlnArgLeuIleAsnAsnAsnTrrPgluPheArgProLysArgLeuAsnPheLysLeuPhe 180

Db 892 CAGCGACTCATCAACAACAATTTGGGATTTCCGGCCCCAAGAGACTCAACTTCAAACCTTTC 951

Qy 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200

Db 952 AACATCCAAGTCAAGGAGGTCAACGCAATGATGGCGTCACAACCATCGCTAATAACCTT 1011

Qy 201 ThrSerThrValGlnValPheSerAspSerGluTrrPgluLeuProTrrValLeuGlySer 220

Db 1012 ACCAGCACGGTTCAAGTCTTCTCGGACTCGGAGTACAGACTTCCGTACGTCTCTCGGCTCT 1071

Qy 221 AlaHisGlnGlyCysLeuProPropheProAlaAspValPheMetIleProGlnTrrGly 240

Db 1072 GCGCACCGGGCTGCCCTCCCTCCGTTCCCGCGGACGTGTTCATGATTCGCAATACGGC 1131

Qy 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTrrCysLeuGlu 260

Db 1132 TACCTGACGCTCAACAATGGCAGCCCAAGCCGCTGGGACGTTTCATCTTTTACTGCCTGGA 1191









QY 221 AlaHisGlnGlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGly 240  
|||  
Db 3279 GCGACACAGGGCTGCCCTCCCTCCGTCGGCGGAGCTGTTTCATGATTCGCGAGTACGGC 3338  
  
QY 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260  
|||  
Db 3339 TACCTAACGCTCAACAATGGCAGCCAGGCGAGTGGGAGTCCCTTTTACTGCTGGAA 3398  
  
QY 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280  
|||  
Db 3399 TATTTCCTATCGCAGATGCTGAGAACGGGCAATAACTTTACCTTCAGCTACACCTTCGAG 3458  
  
QY 281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300  
:::  
Db 3459 GACGTGCCTTTCCACAGCAGCTACGCGCACAGCCAGAGCCTGGACCGGTGATGAATCCT 3518  
  
QY 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320  
|||  
Db 3519 CTCATCGACCATCTGTATTACCTGAACAGAACTCAGAATCAGTCCGGGAAGTCCCCAA 3578  
  
QY 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340  
|||  
Db 3579 AACAAAGGACTTGCTGTTAGCCGGGGGTCTCCAGCTGGCATGTCTGTTTACGCCCAAAAC 3638  
  
QY 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360  
|||  
Db 3639 TGGCTACCTGGACCCCTGTACCAGGACAGCGCGTCTTCAAAATATAAACCCTTAATGGCGTGAATCTATA 3698  
  
QY 361 AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380  
|||  
Db 3699 AACAGCAACTTTACCTGGAGTGGTGTCTTCAAAATATAAACCCTTAATGGCGTGAATCTATA 3758  
  
QY 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400  
|||  
Db 3759 ATCAACCCCTGGCACTGCTATGGCCTCACACAAAGACACAAAGACAAAGTCTTCTTCCCATG 3818  
  
QY 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420  
|||  
Db 3819 AGCGGTGTCATGATTTTGGAAAGGAGAGCGCGGAGCTTCAACACACTGCAATGGACAAT 3878  
  
QY 421 ValMetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440  
|||  
Db 3879 GTCATGATCACAGACGAGAGGAAATCAAAAGCCACTAACCCCGTGGCCACCGAAGATTT 3938  
  
QY 441 GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis 460  
|||  
Db 3939 GGGACTGTGGCAGTCAATCTCAGAGCAGCAGCAGACACCCCTGCGACCGGAGATGTGCAT 3998  
  
QY 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 480  
|||  
Db 3999 GTTATGGAGCCTTACCTGGAATGGTGTGGCAAGACAGACAGACGTATACCTGCGGGTCTCT 4058  
  
QY 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500  
|||  
Db 4059 ATTTGGGCCAAAATTCCTCACACGGATGGACACTTTCACCCGTCTCCTCTCATGGCGGC 4118  
  
QY 501 PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520  
:::  
Db 4119 TTGGACTTAAGCACCCCGCTCTCAGATCCTCATCAAAACACGCTGTTCTCTCGGAAT 4178  
  
QY 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540  
|||  
Db 4179 CCTCCGGCAGAGTTTTCGGCTACAAAGTTTGTCTTCATTCATCACCAGTATTCACAGGA 4238  
  
QY 541 GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro 560  
|||  
Db 4239 CAAGTGACGCTGAGATTGAATGGGAGCTGCAGAAAGAAACAGCAACCGCTGGAATCCC 4298  
  
QY 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580  
|||  
Db 4299 GAAGTGACGATATACATCTAACTATGCAAAATCTGCCAACGTTGATTTTCACTGTGGACAAC 4358

QY 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599  
|||  
Db 4359 AATGGACTTATACTGAGCCTCGCCCATTTGGCACCCTTACCTACCCGTCCCTG 4415  
  
RESULT 5  
AAD00779  
ID AAD00779 standard; DNA; 1605 BP.  
XX  
AC AAD00779;  
XX 08-SEP-2000 (first entry)  
XX Adeno-associated virus serotype 1 capsid protein VP3 DNA.  
DE Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;  
KW cap protein; recombinant viral vector; gene delivery; gene therapy;  
KW vaccine; transgene; VP3; ss.  
XX Adeno associated virus serotype 1.  
OS  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1605  
FT /\*tag= a  
FT /product= "VP3 protein"  
XX  
PN WO200028061-A2.  
XX 18-MAY-2000.  
XX 02-NOV-1999; 99WO-US25694.  
XX 05-NOV-1998; 98US-0107114.  
XX (UYPE-) UNIV PENNSYLVANIA.  
XX Wilson JM, Xiao W;  
XX WPI; 2000-376571/32.  
DR P-PSDB; AAY71169.  
XX  
PT Novel adeno-associated virus serotype 1 polynucleotide useful for  
XX preparation of medicament for delivery of a transgene to a host -  
PS Claim 10; Page 96-99; 108pp; English.  
XX  
CC The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA  
CC which is characterised by two inverted terminal repeats (ITR) and open  
CC reading frames for rep and capsid (cap) proteins. The rep reading frame  
CC encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap  
CC reading frame encodes three structural proteins, VP1, VP2 and VP3.  
CC The AAV-1 sequence or its fragments particularly ITRs, rep and cap  
CC coding regions, are useful in production of recombinant viral vectors  
CC for gene delivery. These vectors can be used as gene therapy  
CC vectors, vaccine vectors or antisense delivery vectors. The AAV-1 does  
CC not induce the formation of neutralising antibodies specific to any  
CC serotype of AAV hence is useful for transforming host cells, and in the  
CC preparation of a medicament for the delivery of transgene to a host.  
CC The present sequence is an AAV-1 DNA encoding a cap protein VP3 which is  
CC useful in the production of recombinant viral vector for gene delivery.  
XX  
SQ Sequence 1605 BP; 431 A; 466 C; 361 G; 347 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 3.64e-231 Length: 1605  
Score: 2906.00 Matches: 534  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 89.39% Indels: 0  
DB: 21 Gaps: 0  
  
US-09-807-802A-15 (1-599) x AAD00779 (1-1605)  
QY 66 MetAlaSerGlyGlyAlaProMetAlaAspAsnGluGlyAlaAspGlyValGly 85



|||||  
1 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCCGACGGAGTGGT 60  
QY AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 105  
Db AATGCCCTCAGGAAATTTGGCATTTGCCATTCACATGGCTGGGCACAGAGTCATCACACC 120  
QY SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 125  
Db AGCACCCGACCTGGGCCTTGCCCACTTACCAATAACCACTCTACAAAGCAATCTCCAGT 180  
QY AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 145  
Db GCTTCAACGGGGCCAGCAACGACCACTACTTCGGCTACAGCACCCCTGGGGGTAT 240  
QY PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 165  
Db TTTGATTTCAACAGATTCCACTGCCACTTTTACCACCGTACTGGCAGCGACTCATCAAC 300  
QY AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 185  
Db AACAAATTGGGATTCCGGCCCCAAGAGACTCAACTTCAAACCTTCAACATCCAAGTCAAG 360  
QY GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 205  
Db GAGGTACGACGAATGATGGCGTCAACAACCATCGCTAATAACCTTACCAGCACGGTTCAA 420  
QY ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 225  
Db GTCTTCTCGGACTCGGAGTACCAGCTTCGGTACGTCTCGGCTCTGGCACCAGGGCTGC 480  
QY LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 245  
Db CTCCCTCCGTTCCGGCGGACGTTTCATGATTCGCAATACGGCTACCTGACGCTCAAC 540  
QY AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 265  
Db AATGGCAGCCCAAGCCGTGGACGTTTCATCTCTTTACTGCCCTGGAATATTCCCTTCTCAG 600  
QY MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPropheHis 285  
Db ATGCTGAGAACGGGCAACAACCTTTACCTTCAGCTACACCTTTGAGGAAGTGCTTTCCAC 660  
QY SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 305  
Db AGCAGCTACGGCACAGCCAGACGCTGGACCGGCTGATGAATCTCTCATCGACCAATAC 720  
QY LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 325  
Db CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAAGGACTTGCTG 780  
QY PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 345  
Db TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACTGGCTACCTGACCC 840  
QY CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 365  
Db TGTATCGGCAGACGCGGTTTCTAAACAAAAACAGACAACAACAACAGCAATTTTACC 900  
QY TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 385  
Db TGGACTGGTGTCTTCAAAATATAACCTCAATGGGCGTGAATCCATCATCAACCCCTGGCACT 960  
QY AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 405  
Db GCTATGGCTCACACAAAGACGACGAAGACAGTCTTTCCCATGAGCGGTGTCATGATT 1020  
QY PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 425  
Db TTTGGAAAAGAGAGCGCGGAGCTTCAAACACTGCATTTGGACAATGTGATGATTACAGAC 1080  
QY GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 445

Db 1081 GAAGAGGAAATTAAAGCCACTAACCTGTGTGGCCACCGAAAGATTTTGGGACCGTGGCAGTC 1140  
QY AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 465  
Db AATTTCCAGAGCAGCAGACAGACCCCTGCGACCGGAGATGTGCATGTATGGAGCATTA 1200  
QY ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 485  
Db CCTGGCATGGTGTGGCAAGATAGAGACGTCGTACCTGCAGGGTCCCATTTTGGGCCAAAATT 1260  
QY ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 505  
Db CCTCACACAGATGGACACTTTCACCCGTCCTCTTATGGGCGGCTTTGGACTCAAGAAC 1320  
QY ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 525  
Db CCGCTCTCAGATCCTCATCAAAACACGCTGTCTCGAATCCTCCGGCGGAGTTT 1380  
QY SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 545  
Db TCAGCTACAAAGTTTGTTTCATTTCATCACCCTACTCCACAGGACAAGTGAGTGTGGAA 1440  
QY IleGluTrpGluLeuGlnLysGluAsnSerLysArgTyrAsnProGluValGlnTyrThr 565  
Db ATTGAATGGGAGCTGCAGAAAGAAACAGCAAGCCGTGGAATCCCGAAGTGCAGTACACA 1500  
QY SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 585  
Db TCCAATTATGCAAAATCTGCCAACGTTTGATTTTACTGTGGACAACAAATGGACTTTTACT 1560  
QY GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599  
Db GAGCCTCGCCCCATTGGCACCCGTTACCTTACCCGTCCTCTG 1602

RESULT 6

AAF23748  
ID AAF23748 standard; DNA; 4722 BP.  
XX  
AC AAF23748;  
XX 28-MAR-2001 (first entry)  
DT  
XX AAV3B DNA sequence.  
DE  
KW AAV3B; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;  
KW atherosclerosis; sickle cell anaemia; thalassaemia;  
KW blood clotting disorder; diabetes; ss.  
OS Adeno associated virus.  
XX  
PN US6156303-A.  
XX  
PD 05-DEC-2000.  
XX  
PF 11-JUN-1997; 97US-0873168.  
XX  
PR 11-JUN-1997; 97US-0873168.  
XX (UNIW ) UNIV WASHINGTON.  
PA  
XX Russell DW, Rutledge EA;  
PI  
XX WPI; 2001-060164/07.  
DR  
XX Adeno-associated virus serotype 6 and viral vector derived from it for  
PT gene therapy of cystic fibrosis, cancer, acquired immunodeficiency  
PT syndrome, sickle cell anemia, thalassemia and diabetes -  
XX  
PS Example 2; Fig 1; 50pp; English.  
XX  
CC The present invention relates to adeno-associated virus serotypes. The  
CC present sequence is the DNA sequence of one such serotype (AAV3B). AAV3B  
CC can be used to construct AAV viral vectors for use in gene therapy for a

CC range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,  
CC sickle cell anaemia, thalassaemia, blood clotting disorders and diabetes.  
CC The AAV viral vectors have increased transduction efficiency of a  
CC particular host cell as the AAV virion containing the AAV vector genome  
CC can be modified to express a capsid protein of an AAV serotype that  
CC transduces the selected host cell.  
XX  
SQ Sequence 4722 BP; 1204 A; 1297 C; 1243 G; 978 T; 0 other;

Alignment Scores:  
Pred. No.: 2.13e-224 Length: 4722  
Score: 2832.00 Matches: 514  
Percent Similarity: 91.33% Conservative: 34  
Best Local Similarity: 85.67% Mismatches: 50  
Query Match: 87.11% Indels: 2  
DB: 22 Gaps: 2

US-09-807-802A-15 (1-599) x AAF23748 (1-4722)

QY 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20  
Db 2619 ACGGCTCCTGGAAAGAGAGGCCTGTAGATCAGTCTCTCAGGAACCGGACTCATCT 2678  
QY 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40  
Db 2679 GGTGTTGGCAATCGGGCAACAGCCTGCCAGAAAAAGACTAAATTTTCGGTCAGACTGGC 2738  
QY 41 AspSerGluSerValProAppProGlnProLeuGlyGluProProAlaThrProAlaAla 60  
Db 2739 GACTCAGAGTCAGTCCCAGACCCTCAACCTCTCGGAGAACCCAGCAGCCCCCAAGT 2798  
QY 61 ValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGly 80  
Db 2799 TTGGGATCTAATAACAATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATACGAGGT 2858  
QY 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100  
Db 2859 GCCGATGGAGTGGTAAATTCCTCAGGAAATTGGCAATGGCATTCCTCAATGGCTGGCGGAC 2918  
QY 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120  
Db 2919 AGAGTCATCACCACCAGCACCAACCTGGGCCCTGCCACTTACCAACAACCATCTCTAC 2978  
QY 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140  
Db 2979 AAGCAAAATCTCCAGCCAATCA--GGAGCTTCAAACGACAACCACTACTTTGGCTACAGC 3035  
QY 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160  
Db 3036 ACCCTTGGGGGTATTTTGACTTTAACAGATTCCACATGCCACTTCTCACCACCGTACTGG 3095  
QY 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180  
Db 3096 CAGCGACTCATTAACAACAACCTGGGGATTCCGGCCCCAAGAACTCAGCTTCAAGCTCTTC 3155  
QY 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200  
Db 3156 AACATCCAAGTTAAAGAGGTACGCAGAACGATGGCAGCAGACTATTGCCAATAACCTT 3215  
QY 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220  
Db 3216 ACCAGCAGGTTCAAGTGTTCAGGACTCGGAGTATCAGTCCCGTACGTCTCGGGTTCG 3275  
QY 221 AlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGly 240  
Db 3276 GCGCACCAAGGCTGTCTCCCGCGGTTCAGCGGACGCTTTCATGGTCCCTCAGTATGGA 3335  
QY 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260  
Db 3336 TACCTCACCTTGAACAACGGAAGTCAAGCGGTGGGACGCTCATCTTTTACTGCTGGAG 3395  
QY 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280

Db 3396 TACTTCCCTTCGCAGATGCTAAGGACTGGAAATAACTTCCAATTTCAGCTATACCTTCGAG 3455  
QY 281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300  
Db 3456 GATGTACTTTTTCACAGCAGCTACGCTCACAGCCAGAGTTTGGATCGCTTGATGAATCCT 3515  
QY 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGln---AsnGlnSerGlySerAla 319  
Db 3516 CTTATTGATCAGTATCTGTACTACTGAAACAGAACGCAAGGAACAACCTCTGGAACAAC 3575  
QY 320 GlnAsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLys 339  
Db 3576 AACCAATCAGGCTGCTTTTTCAGCAGGCTGGCCTCAGTCTATGTCTTTCGAGGCCAGA 3635  
QY 340 AsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsn 359  
Db 3636 AATTGGCTACCTGGGCCCTGCTACCGGCAACAGAGACTTTTCAAAGACTGCTAAACGACAAC 3695  
QY 360 AsnAsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSer 379  
Db 3696 AACAAACAGTAACCTTCTTGGACAGCGGCCAGCAAAATATCATCTCAATGGCCGACTCG 3755  
QY 380 IleIleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePhePro 399  
Db 3756 CTGGTGAATCCAGGACCAGCTATGGCCAGTCAACAGGACGATGAAGAAAAATTTTCCCT 3815  
QY 400 MetSerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAsp 419  
Db 3816 ATGCACGGCAATCTAATATTTGGCAAGAAGGACAAACGGCAAGTAACGCAGAATTAGAT 3875  
QY 420 AsnValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArg 439  
Db 3876 AATGTAATGATTACGGATGAAGAGAGATTTCGTACCAATCTCTGTGGCAACAGAGCAG 3935  
QY 440 PheGlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspVal 459  
Db 3936 TATGGAATCTGGCAATAAATTTGCAGAGCTCAAATACAGCTCCCACGACTAGAACTGTC 3995  
QY 460 HisAlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGly 479  
Db 3996 AATGATCAGGGGCTTACCTGGCATGGTGGCAAGATCGTGACGTGTACCTTCAAGGA 4055  
QY 480 ProIleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGly 499  
Db 4056 CCTATCTGGCAAGATTCTCTCACCGGATGGACACTTTCATCTCTCTCTCTGTATGGA 4115  
QY 500 GlyPheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAla 519  
Db 4116 GGCTTTGGACTGAAACATCCGCCCTCTCAATCATGATCAAAAATACTCCGGTACCGGCA 4175  
QY 520 AsnProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThr 539  
Db 4176 AATCCTCCGACGACTTTCAGCCCCGCAAGTTTGTCTTCAATTTATCACTCAGTACTCCACT 4235  
QY 540 GlyGlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsn 559  
Db 4236 GGACAGGTCAGCGTGGAAATTTGAGTGGGAGCTACAGAAAGAAAAACAGCAAAACGTTGGAA 4295  
QY 560 ProGluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAsp 579  
Db 4296 CCAGAGATTCAGTACACTTCCAACACTACAACAAGTCTGTTAATGTGGACTTTACTGTAGAC 4355  
QY 580 AsnAsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599  
Db 4356 ACTAATGGTGTATTAGTGAACCTCGCCCTATTGGAAACCCCGGTATCTCACACGAAACTTG 4415

RESULT 7

AAD00834

ID AAD00834 standard; DNA; 8178 BP.

XX AAD00834;

DT 08-SEP-2000 (first entry)

XX DE Chimeric adeno-associated virus 2/3 capsid encoding helper plasmid.

XX KW Adeno-associated virus; AAV2; AAV3; chimeric; gene therapy; plasmid;

KW KW recombinant parvoviral vector; cellular tropism; rep gene; cap protein;

XX KW capsid; gene delivery; VP1; VP2; VP3; ss.

XX OS Chimeric - Adeno associated virus serotype 2.

OS OS Chimeric - Adeno associated virus serotype 3.

XX FH Key Location/Qualifiers

FT CDS 251..2180

FT /\*tag= a

FT /product= "Rep 68"

FT /note= "The specification states that this region

FT encodes Rep 68 protein of AAV2. However, the stop

FT position given in the specification for this CDS does

FT not correspond to stop codon"

FT CDS 251..2114

FT /\*tag= b

FT /product= "Rep 78"

FT /note= "The specification states that this region

FT encodes Rep 78 protein of AAV2. However, the stop

FT position given in the specification for this CDS does

FT not correspond to stop codon"

FT CDS 923..2114

FT /\*tag= c

FT /product= "Rep 52"

FT /note= "The specification states that this region

FT encodes Rep 52 protein of AAV2. However, the stop

FT position given in the specification for this CDS does

FT not correspond to stop codon"

FT CDS 923..2180

FT /\*tag= d

FT /product= "Rep 40"

FT /note= "The specification states that this region

FT encodes Rep 40 protein of AAV2. However, the stop

FT position given in the specification for this CDS does

FT not correspond to stop codon"

FT CDS 2133..4342

FT /\*tag= e

FT /product= "VP1 cap protein"

FT /note= "The specification states that this region

FT encodes VP1 protein of AAV2. However, the stop

FT position given in the specification for this CDS does

FT not correspond to stop codon"

FT CDS 2544..4342

FT /\*tag= f

FT /product= "VP2 cap protein"

FT /note= "The specification states that this region

FT encodes VP2 protein of AAV2. However, the positions

FT given in the specification for this CDS do not correspond

FT to start or stop codons"

FT CDS 2739..4342

FT /\*tag= g

FT /product= "VP3 cap protein"

FT /note= "The specification states that this region

FT encodes VP3 protein of AAV2 in which loops 2-4 are

FT replaced with the corresponding region from AAV3.

FT However, the stop position given in the specification

FT for this CDS does not correspond to stop codon"

FT misc\_feature 3184..4092

FT /\*tag= h

FT /note= "AAV3 sequence"

FT XX WO200028004-A1.

XX PD 18-MAY-2000.

XX PF 10-NOV-1999; 99WO-US26505.

XX PR 10-NOV-1998; 98US-0107840.

PR 10-MAR-1999; 99US-0123651.

XX PA (UYNC-) UNIV NORTH CAROLINA.

XX PI Rabinowitz JE, Samulski RJ, Xiao W;

XX DR WPI; 2000-376523/32.

XX PT Recombinant parvoviral vectors with altered packaging, tropisms and

XX PT immunogenic properties, useful in gene therapy protocols -

XX PS Example 27; Page 143-147; 153pp; English.

XX CC The patent discloses modified parvovirus vectors with advantageous

CC antigenic properties, packaging capabilities and cellular tropisms.

CC These vectors can be used in standard recombinant DNA protocols e.g. gene

CC therapy for delivering nucleic acids to cells. The present sequence is

CC a helper plasmid encoding a chimeric adeno-associated virus serotype 2/3

CC (AAV2/3) capsid. The sequence contains the AAV2 rep coding sequences,

CC most of the AAV2 capsid coding sequences with the exceptions that

CC loops 2-4 from the AAV2 VP3 subunit is replaced with the corresponding

CC region from AAV3, in a pBluescript backbone. The rep

CC coding sequence encodes four proteins responsible for replication,

CC Rep 68, Rep 78, Rep 52 and Rep 40, and the capsid (cap) coding region

CC encodes three structural proteins VP1, VP2 and VP3.

CC Recombinant parvovirus comprising the chimeric capsid is useful

CC for gene delivery.

XX SQ Sequence 8178 BP; 2082 A; 2078 C; 2027 G; 1991 T; 0 other;

Alignment Scores:

Pred. No.: 1.66e-221 Length: 8178

Score: 2801.00 Matches: 509

Percent Similarity: 90.33% Conservative: 33

Best Local Similarity: 84.83% Mismatches: 56

Query Match: 86.16% Indels: 2

DB: 21 Gaps: 2

US-09-807-802A-15 (1-599) x AAD000834 (1-8178)

QY 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20

Db 2544 ACGGCTCCGGGAAAAGAGAGCGGTAGAGCACTCTCTCTGGAGCCAGACTCCTCTCG 2603

QY 21 GlyTleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40

Db 2604 GGAACCGAAAGCGGGGCCAGCAGCCTGCAAGAAAAGATTGAATTTGGTCAGACTGGA 2663

QY 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60

Db 2664 GACGCAGACTCAGTACTGACCCCGAGCCTCTCGGACAGCCAGCCCTCTCTGGT 2723

QY 61 ValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGly 80

Db 2724 CTGGGAACATAATACGATGGCTACAGGCAGTGCGGCACCAATGGCAGACAATAACGAGGC 2783

QY 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100

Db 2784 GCCGACGGAGTGGGTAATTCCTCGGAAATTTGGCATTTGGATTCCACATGGATGGCGGAC 2843

QY 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTrpAsnAsnHisLeuTyr 120

Db 2844 AGAGTCATCACCACCAGCACCCGAACTGGGCCCTGCCCCACCTACAAACACCCTCTAC 2903

QY 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140

Db 2904 AAACAAATTTCCAGCCAATCA---GGAGCCTCGAACGACATCACTACTTTGGCTACAGC 2960

QY 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160

Db 2961 ACCCCTTGGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCACCACGTGCTGG 3020

QY 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180



Db 3021 CAAGACTCATCAACAACAACCTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTT 3080

Qy 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200

Db 3081 AACATTCAAGTCAAAGAGGTACGCAGAAATGACGGTACGACGACGATTGCCAATAACCTT 3140

Qy 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220

Db 3141 ACCAGCAGGTTTCAGGTGTTTACTGACTCGGAGTACCACTCCCTACGTCTCGGTGCG 3200

Qy 221 AlaHisGlnGlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGly 240

Db 3201 GCGCAACAAGGCTGTCTCCCGCGCTTCCAGCGGACGTCTTCATGCTCCCTCAGTATGGA 3260

Qy 241 TyrLeuThrLeuAsnAsnGlnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260

Db 3261 TACCTCACCTTGAACAACGGAAGTCAAGCGGTGGGACGCTCATCTTTTACTGCCTGGAG 3320

Qy 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280

Db 3321 TACTTCCCTTCGCAGATGCTAAGGACTGGAAATAACTTCCAATTCACTATACCTTCGAG 3380

Qy 281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300

Db 3381 GATGTACCTTTTCACAGCAGCTACGCTCACAGCCAGAGTTTGGATCGCTTGATGAATCCT 3440

Qy 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGln---AsnGlnSerGlySerAla 319

Db 3441 CTTATTGATCAGTATCTGTACTACTGAACAGAACGCAAGCAACAACCTCTGGAAACAAC 3500

Qy 320 GlnAsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLys 339

Db 3501 AACCAAATCACGGCTGCTTTTATGCCAGGCTGGCCCTCAGTCTATGCTTTTGCAGGCCAGA 3560

Qy 340 AsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsn 359

Db 3561 AATTGGCTACCTGGGCCCTGCTACCGGCAACAGAGACTTTCAAAGACTGCTAACGACAAC 3620

Qy 360 AsnAsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSer 379

Db 3621 AACAACAGTAACTTTCTTGGACAGCGGCCAGCAAAATATCATCTCAATGGCCGCGACTCG 3680

Qy 380 IleIleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePhePro 399

Db 3681 CTGGTGAATCCAGGACCAGCTATGGCCAGTCACAAGGACGATGAAGAAAATTTTTCCT 3740

Qy 400 MetSerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAsp 419

Db 3741 ATGCACGGCAATCTAATATTTGGCAAGAAAGGGCAACCGCAAGTAACGCAGAAATTAGAT 3800

Qy 420 AsnValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArg 439

Db 3801 AATGTAATGATTACGGATGAAGAGAGATTTCGTACCACCAATCTGTGGCAACAGAGCAG 3860

Qy 440 PheGlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspVal 459

Db 3861 TATGGAACTGTGGCAAAATAACTTGCAGAGCTCAAATACAGCTCCACGACTGGAACTGTC 3920

Qy 460 HisAlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGly 479

Db 3921 AATCATCAGGGGGCCTTACCTGGCATGGTGTGGCAAGATCGTGACGTGTACCTTCAAGGA 3980

Qy 480 ProIleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGly 499

Db 3981 CCTATCTGGGCAAGATTCTCTCACCGGATGGACACTTTTCATCTCTCTCTCTGATGGGA 4040

Qy 500 GlyPheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAla 519

Db 4041 GGCTTTGGACTGAACAACATCCGCCCTCCTCAAATCATGATCAAAAATACTCCGGTACTCGG 4100

Qy 520 AsnProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThr 539

Db 4101 AATCCTTCGACCACCTTCAGTGGCGCAAGAGTTTGCTTCTTCATCATCACAGTACTCCAGG 4160

Qy 540 GlyGlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsn 559

Db 4161 GGACAGGTCAGCGTGGAGATCGAGTGGAGCTGCAGAAAGGAAACAGCAACGCTGGAAT 4220

Qy 560 ProGluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAsp 579

Db 4221 CCCGAAATTCAGTACACTTCCAACTACAAAGTCTGTTAATCGTGGACTTACCGTGGAT 4280

Qy 580 AsnAsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599

Db 4281 ACTAATGGCGTGTATTTCAGAGCCCTCGCCCATTCGCCACAGATACCTGACTCGTAATCTG 4340

RESULT 8

AAH26327

ID AAH26327 standard; DNA; 4072 BP.

XX

AC AAH26327;

XX

DT 02-OCT-2001 (first entry)

XX

DE Adenovirus helper Ad cap2.

XX

KW AAV; vector; adenovirus; helper virus; Ad cap2; loxp site;

KW gene therapy; ds.

XX

OS Chimeric - Mastadenovirus.

OS Chimeric - Adeno associated virus.

OS Chimeric - Human cytomegalovirus.

XX

PN WO200155361-A2.

XX

PD 02-AUG-2001.

XX

PF 26-JAN-2001; 2001WO-US02709.

XX

PR 26-JAN-2000; 2000US-0178536.

XX

PA (CHIR ) CHIRON CORP.

XX

PI Hardy SP;

XX

DR WPI; 2001-483239/52.

XX

PT Producing recombinant adeno-associated virus (rAAV) vector, by stably

PT infecting eukaryotic host cell with rAAV vectors, helper adenovirus or

PT herpes virus, recombinant adenovirus or herpes vectors -

XX

PS Example 2; Page 58-59; 63pp; English.

XX

CC The present sequence is that of adenovirus helper Ad cap2, made

CC by Cre/lox recombination and expressing adeno-associated virus

CC (AAV) VPI, 2,3 from the cytomegalovirus immediate early promoter

CC of padlox, and containing a loxp site. Ad cap2 was used to

CC demonstrate recombinant AAV (rAAV) production from virally

CC transduced cells. The invention provides methods and compositions

CC for producing rAAV vector particles by: (a) introducing into a host

CC cell (i) AAV packaging plasmid pfloxAAV (see AAH26326), (ii) a

CC recombinant viral vector encoding plasmid, and (iii) a plasmid

CC encoding herpesvirus, cytomegalovirus or adenoviral functions, or

CC a herpesvirus, cytomegalovirus or adenovirus itself, to produce

CC flox AAV particles and rAAV particles; and (b) introducing into a

CC second host cell (i) the rAAV particles or (a), (ii) a vector

CC that directs expression of Cre, and (ii) a vector which directs

CC expression of herpes virus, cytomegalovirus or adenovirus helper

CC functions, such that rAAV vector particles are produced. The

CC vectors are useful for in vivo or in vitro gene therapy and also

CC for in vitro recombinant protein production.

SQ Sequence 4072 BP; 1073 A; 1025 C; 992 G; 982 T; 0 other;

Alignment Scores:

Pred. No.: 6.84e-219 Length: 4072

|   |         |   |      |
|---|---------|---|------|
| Score:  | 2764.50 | Matches:  | 496  |
| Percent Similarity:                           | 89.98%  | Conservative:   | 43   |
| Best Local Similarity:                        | 82.80%  | Mismatches:   | 59   |
| Query Match:                                  | 85.04%  | Indels:   | 1    |
| DB:   | 22      | Gaps:   | 1    |
| US-09-807-802A-15 (1-599) x AAH26327 (1-4072) |         |   |      |
| QY  | 1       | ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer  | 20   |
| Db  | 1895    | ACGGCTCCGGGAAAAAGAGGCCGTAGAGCACTCTCCTGTGGAGCCAGACTCCTCTCTCG   | 1954 |
| QY  | 21      | GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly  | 40   |
| Db  | 1955    | GGAAACCGGAAAGCGGGCCAGCAGCTGCAAGAAAAAGATTGAATTTTGGTCAGACTGGA   | 2014 |
| QY  | 41      | AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla  | 60   |
| Db  | 2015    | GACGCAGACTCAGTACCTGACCCCGAGCTCTCGGACAGCCAGCAGCCCTCTGGT        | 2074 |
| QY  | 61      | ValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnGluGly     | 80   |
| Db  | 2075    | CTGGGAACATAATACGATGGCTACAGGCAGTGGCGCACCAATGCGAGACAATAACGAGGSC | 2134 |
| QY  | 81      | AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp  | 100  |
| Db  | 2135    | GCCGACGGAGTGGTAATTCTCTCGGGAATTGGCATTGCGATTCCACATGGATGGGCGAC   | 2194 |
| QY  | 101     | ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTrpAsnAsnHisLeuTyr  | 120  |
| Db  | 2195    | AGAGTCATCACCCAGCACCCGAACTGGGCCCTGCCACCTACAACAACCACTCTAC       | 2254 |
| QY  | 121     | LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer  | 140  |
| Db  | 2255    | AAACAAATTTCCAGCCAATCA---GGAGCCTCGAACGACAATCACTACTTTGGCTACAGC  | 2311 |
| QY  | 141     | ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp  | 160  |
| Db  | 2312    | ACCCCTTGGGGTATTTGACTTCAACAGATTCCACTGCCACTTTTCCACCACGTGACTGG   | 2371 |
| QY  | 161     | GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe  | 180  |
| Db  | 2372    | CAAGACTCATCAACAACAACCTGGGGATTCCGAGCCCAAGAGACTCAACTTCAAGCTCTT  | 2431 |
| QY  | 181     | AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu  | 200  |
| Db  | 2432    | AACATTCAAGTCAAGAGGTCAACGAGAAATGACGGTACGACGACGATTGCCAATAACCTT  | 2491 |
| QY  | 201     | ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer  | 220  |
| Db  | 2492    | ACCAGCACGGTTCAGGTGTTTACTGACTCGAGTACCAGCTCCCGTACGTCCTCGGCTCG   | 2551 |
| QY  | 221     | AlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGly  | 240  |
| Db  | 2552    | GCGCATCAAGGATGCTCTCCGCCGTTCACGAGACAGCTCTTTCATGGTGCACAGTATGGA  | 2611 |
| QY  | 241     | TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu  | 260  |
| Db  | 2612    | TACCTCACCTGAAACAACGGGAGTCAGGCAGTAGGACGCTCTTCAATTTACTGCTGGAG   | 2671 |
| QY  | 261     | TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu  | 280  |
| Db  | 2672    | TACTTTCTCTCAGATGCTGCGTACCGGAAACAACCTTACCTTCAGCTACACTTTTGAG    | 2731 |
| QY  | 281     | GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro  | 300  |
| Db  | 2732    | GACGTTCTTTCCACAGCAGCTACGCTCAAGCCAGAGTCTGGACCTCTCATGAATCCT     | 2791 |
| QY  | 301     | LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln  | 320  |
| Db  | 2792    | CTCATCGACCACTGCTGTTACTTTGAGCAGACAACAACACTCCCAAGTGGAAACCCACG   | 2851 |
| QY  | 321     | AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn  | 340  |

|          |  |   |      |
|----------|--|---|------|
| Db       | 2852   | CAGTCAAGGCTTCAGTTTTTCTCAGGCCCGGACCGAGTGACATTCGGGACCACTTAGGAAC   | 2911 |
| QY       | 341  | TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn    | 360  |
| Db       | 2912   | TGGCTTCTCGGACCTGTATTACCGCCAGCAGCAGATATCAAAGACATCTGCGGATAACAAC   | 2971 |
| QY       | 361  | AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle    | 380  |
| Db       | 2972   | AACAGTGAATACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTG    | 3031 |
| QY       | 381  | IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet    | 400  |
| Db       | 3032   | GTGAATCCGGGCCCGCCATGGCAAGCCACACAGGACGATGAAGAAAAGTTTTCCTCAG      | 3091 |
| QY       | 401  | SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn    | 420  |
| Db       | 3092   | AGCGGGTTCTCATCTTTGGGAAGCAAGGCTCAGAGAAAAACAATGTGGACATTTGAAAAAG   | 3151 |
| QY       | 421  | ValMetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe    | 440  |
| Db       | 3152   | GTCAATGATTACAGACGAAGGAAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTAT     | 3211 |
| QY       | 441  | GlyThrValAlaValAsnPheGlnSerSerSerSerThrAspProAlaThrGlyAspValHis | 460  |
| Db       | 3212   | GGTTCTGTATCTACCAACCTCCAGAGAGGGCAACAGACAGCTACCCGACAGATGTCAAC     | 3271 |
| QY       | 461  | AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro    | 480  |
| Db       | 3272   | ACACAAGCGGTTCTTCCAGGCATGGTCTGGCAGACAGAGATGTGTACTCTCAGGGGCCC     | 3331 |
| QY       | 481  | IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly    | 500  |
| Db       | 3332   | ATCTGGGCAAGATTCCACACACGAGCGGACATTTTCACCCCTCTCCCTCATGGGTGGA      | 3391 |
| QY       | 501  | PheGlyLeuLysAsnProProGlnIleLeuLysAsnThrProValProAlaAsn          | 520  |
| Db       | 3392   | TTCCGACTTAAACACCCCTCTCCACAGATTCTCATCAAGAAACACCCCGGTACCTGCGAAT   | 3451 |
| QY       | 521  | ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly    | 540  |
| Db       | 3452   | CCTTCGACCACTTTCAGTCGGGCAAGTTTGTCTTCTCATCACACAGTACTCCACGGGA      | 3511 |
| QY       | 541  | GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro    | 560  |
| Db       | 3512   | CAGTCAAGCTGGAGATCGAGTGGGAGCTGCAGAGGAAACAGCAACGCTGGAATCCC        | 3571 |
| QY       | 561  | GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn    | 580  |
| Db       | 3572   | GAAATTCAGTACACTTCCAACTACAACAGTCTGTAAATGTGGACTTTTACTGTGGACACT    | 3631 |
| QY       | 581  | AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu       | 599  |
| Db       | 3632   | AATGGCGTGTATTTCAGAGCCTCGCCCATTTGGCACCAATACCTGACTCGTAATCTG       | 3688 |
| RESULT 9 |  |   |      |
| ID       | AAI66974   | standard; DNA; 4679 BP.   |      |
| XX       | AAI66974;  |   |      |
| AC       | AAI66974;  |   |      |
| DT       | 11-FEB-2002  | (first entry)   |      |
| XX       | Adeno-associated virus 2   | complete genomic sequence.                                      |      |
| DE       | Recombinant viral vector; RVV; capsid; parvovirus; transgene; cytosstatic; |   |      |
| XX       | inverted terminal repeat; neurotropic; neuroprotective; antianemic; ITR;   |   |      |
| KW       | antidiabetic; antitumour; gene therapy; adeno-associated virus; AAV; ds.   |   |      |
| XX       | Adeno-associated virus 2.  |   |      |
| OS       |  |   |      |
| XX       | Key  | Location/Qualifiers   |      |
| FT       | CDS  | 321..2252   |      |

FT /tag= a  
FT /product= "Rep 68 protein"  
FT /note= "contains introns; for coding sequence  
FT join (321..1906, 2228..2252)"  
FT CDS 321..2186  
FT /tag= b  
FT /product= "Rep 78 protein"  
FT 993..2252  
FT /tag= c  
FT /product= "Rep 40 protein"  
FT /note= "contains introns; for coding sequence  
FT join (993..1906, 2228..2252)"  
FT CDS 993..2186  
FT /tag= d  
FT /product= "Rep 52 protein"  
FT 2203..4410  
FT /tag= e  
FT /product= "major coat protein VP1"  
FT 2614..4410  
FT /tag= f  
FT /transl\_except= "(pos:2614..2616, aa: Met)"  
FT /product= "major coat protein VP2"  
FT 2809..4410  
FT /tag= g  
FT /product= "major coat protein VP3"

WO200168888-A2.

20-SEP-2001.

13-MAR-2001; 2001WO-US07927.

14-MAR-2000; 2000US-189110P.

(NEUR-) NEUROLOGIX INC.

Xiao W, During MJ;

WPI; 2001-596912/67.

P-PSDB; AAG65788, AAG65789, AAG65790, AAG65791, AAG65792, AAG65793, AAG65794.

Recombinant viral vector useful in improving gene therapy in a subject, and for increasing efficiency of entry into a cell, comprises a chimeric capsid having one non-native amino acid sequence and a desired transgene .

Disclosure; Page 46-47; 53pp; English.

The invention provides a recombinant viral vector (RVV) comprising a chimeric capsid (I) having at least one non-native amino acid sequence, derived from a capsid protein domain of parvovirus (II), a virus (III), or their combination, and a transgene flanked 5' and 3' by inverted terminal repeat (ITR) sequences, derived from (II), (III), or their combination. The RVV is useful for improving gene therapy in a subject with a disorder, and for increasing the efficiency of entry into a cell, which involves producing (I) encapsulating a viral vector, and contacting a cell with RVV having (I) such that (I) binds to an attachment site on the cell surface and permits the vector to enter the cell efficiently. A pharmaceutical composition comprising RVV with (I) containing a transgene sequence associated with a disease or a disorder such that expression of the transgene would result in amelioration of the disease or disorder such as inherited neurological and metabolic diseases e.g. lysosomal storage disease, Lesch-Nyhan syndrome, amyloid polynuropathy, Alzheimer amyloid, Duchenne's muscular dystrophy, diseases of the blood e.g. sickle-cell anemia, clotting disorders and thalassemias, cystic fibrosis, diabetes, diseases associated with hormone deficiencies, retinoblastoma and various types of neoplastic cells which include tumours especially central nervous system tumours, neoplasms, carcinomas, sarcomas, leukemias and lymphoma. The present sequence represents the complete genomic sequence of adeno-associated virus 2 which is used in the construction of a chimeric vector.

SQ Sequence 4679 BP; 1198 A; 1262 C; 1255 G; 964 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 8.26e-219 Length: 4679  
Score: 2764.50 Matches: 496  
Percent Similarity: 89.98% Conservative: 43  
Best Local Similarity: 82.80% Mismatches: 59  
Query Match: 85.04% Indels: 1  
DB: 22 Gaps: 1

US-09-807-802A-15 (1-599) x AAI66974 (1-4679)

QY 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20  
Db 2614 ACGGCTCCGGGAAAAAAGAGGCCCGGTAGACACTCTCCTGTGGAGCCAGACTCCTCCTCG 2673  
QY 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40  
Db 2674 GGAACCGGAAAGCGGGCCAGCAGCCTGCAAGAAAAAAGATTGAATTTTGGTCAGACTGGA 2733  
QY 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60  
Db 2734 GACGCAGACTCAGTACCTGACCCCCAGCCTCTCGGACAGCCACCAGCAGCCCTCTTGGT 2793  
QY 61 ValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGly 80  
Db 2794 CTGGGAACATAATACGATGGCTACAGGCAGTGGCGCACCAATGGCAGACATAACGAGGGC 2853  
QY 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100  
Db 2854 GCCGACGGAGTGGGTAATTCCTCGGGAATTTGGCATTCGCATTCACATGGATGGCGAC 2913  
QY 101 ArgValIleThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120  
Db 2914 AGAGTCATCACCCAGCACCCGAAACCTGGGCCCTGCCACCTACACACACCTCTAC 2973  
QY 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140  
Db 2974 AAACAAATTTCCAGCCAATCA--GGAGCCTCGAACGACAATCACTACTTTGGGTACAGC 3030  
QY 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160  
Db 3031 ACCCCTTGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTTCACACCGTACTGG 3090  
QY 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheValLeuPhe 180  
Db 3091 CAAAGACTCATCAACAACAACCTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTT 3150  
QY 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200  
Db 3151 AACATTCAAGTCAAAGAGGTCCAGCAGAAATGACGGTACGACGAGATTGCCAATAACCTT 3210  
QY 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220  
Db 3211 ACCAGCACGGTTCAGGTGTTTACTGACTCGGAGTACCAGTCCCGTACGTCTCTCGGCTCG 3270  
QY 221 AlaHisGlnGlyCysLeuProPropheProAlaAspValPheMetIleProGlnTyrGly 240  
Db 3271 GCGCATCAAGGATGCCTCCCGCCGTTCCAGCAGACGTCTTTCATGGTGCCACAGTATGGA 3330  
QY 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260  
Db 3331 TACCTCACCCCTGAACAACGGGAGTCAGGCAGTAGGACGCTCTTCAATTTTACTGCGCTGAG 3390  
QY 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280  
Db 3391 TACTTTCTCTTCAGATGCTCGTACCGGAAACAACATTACCTTCAGCTACACTTTTGAG 3450  
QY 281 GluValPropheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300  
Db 3451 GACGTTCTCTTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGATCCT 3510  
QY 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320



```
Db      ||| 3511 CTCATCGACCACTAGTATTACTTGAGCAGAAACAACACTCCAAGTGGAAACCAACG 3570
Qy      ||| 321 AsnLysAspLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340
Db      ||| 3571 CAGTCAAGGCTTCAGTTTTCTCAGGCCGAGCGAGTGACATTCGGGACCACTAGGAAC 3630
Qy      ||| 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360
Db      ||| 3631 TGGCTTCCTGGACCCCTGTACCGCCAGCAGCGAGTATCAAAAGACATCTGCGGATAACAAC 3690
Qy      ||| 361 AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380
Db      ||| 3691 AACAGTGAATACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTG 3750
Qy      ||| 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400
Db      ||| 3751 GTGAATCCGGCCCGCCATGSCAAGCCACAAAGGACGATGAAGAAAGTTTTCCTCAG 3810
Qy      ||| 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420
Db      ||| 3811 AGCGGGTTCATCTTTGGGAAGCAAGGCTCAGAGAAACAATGTGGACATTGAAAAG 3870
Qy      ||| 421 ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440
Db      ||| 3871 GTCATGATTACAGACGAAGAGGAATCAGGACCAACCAATCCCGTGGCTACGGAGCAGTAT 3930
Qy      ||| 441 GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis 460
Db      ||| 3931 GGTTCGTATCTACCAACCTCCAGAGAGGCAACAGACAGCAGCTACCGCAGATGTCAAC 3990
Qy      ||| 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 480
Db      ||| 3991 ACACAAGGCGTTCCTCCAGGCATGCTTGGCAGGACAGAGATGTACCTTCAGGGGCC 4050
Qy      ||| 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500
Db      ||| 4051 ATCTGGGCAAGATTCCACACACGAGCGACATTTTCACCCCTCTCCCTCATGSGTGA 4110
Qy      ||| 501 PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520
Db      ||| 4111 TTCGGACTTAACACCCCTCCTCCACAGATTCTCATCAAGAAACACCCCGGTACCTGCGAAT 4170
Qy      ||| 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540
Db      ||| 4171 CCTTCGACCACTTCAGTCCGGCAAGTTTGCTTCCTTCATCACACAGTACTCCACGGGA 4230
Qy      ||| 541 GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro 560
Db      ||| 4231 CAGGTACGGTGGAGATCGAGTGGGAGTCGAGAAGGAAACAGCAACCGTGGAAATCCC 4290
Qy      ||| 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580
Db      ||| 4291 GAAATTCAGTACACTTCCAACACTACAACAAGTCTGTTAATGTGGACTTACTGTGGACACT 4350
Qy      ||| 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599
Db      ||| 4351 AATGGCGTGATTACAGAGCCTCGCCCCATTGGCACCAGATACCTGACTCGTAATCTG 4407
```

RESULT 10

```
AAF23750
ID  AAF23750 standard; DNA; 4679 BP.
XX
AC  AAF23750;
XX
DT  28-MAR-2001 (first entry)
XX
DE  AAV2 DNA sequence.
XX
KW  AAV2; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;
KW  atherosclerosis; sickle cell anaemia; thalassaemia;
KW  blood clotting disorder; diabetes; ss.
```

```
OS      Adeno associated virus.
XX      US6156303-A.
PN
XX      05-DEC-2000.
PD
XX      11-JUN-1997; 97US-0873168.
PF
XX      11-JUN-1997; 97US-0873168.
PR
XX      (UNIW ) UNIV WASHINGTON.
PA
XX      Russell DW, Rutledge EA;
PI
XX      WPI; 2001-060164/07.
DR
XX      Adeno-associated virus serotype 6 and viral vector derived from it for
PT      gene therapy of cystic fibrosis, cancer, acquired immunodeficiency
PT      syndrome, sickle cell anemia, thalassemia and diabetes -
XX      Claim 7; Fig 1; 50pp; English.
PS
XX      The present invention relates to adeno-associated virus serotypes. The
CC      present sequence is the DNA sequence of one such serotype (AAV2). AAV2
CC      can be used to construct AAV viral vectors for use in gene therapy for a
CC      range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,
CC      sickle cell anaemia, thalassaemia, blood clotting disorders and diabetes.
CC      The AAV viral vectors have increased transduction efficiency of a
CC      particular host cell as the AAV virion containing the AAV vector genome
CC      can be modified to express a capsid protein of an AAV serotype that
CC      transduces the selected host cell.
XX
SQ      Sequence 4679 BP; 1197 A; 1263 C; 1255 G; 964 T; 0 other;

Alignment Scores:
Pred. No.:      8.26e-219      Length:      4679
Score:          2764.50      Matches:      496
Percent Similarity: 89.98%      Conservative: 43
Best Local Similarity: 82.80%      Mismatches: 59
Query Match:      85.04%      Indels:      1
DB:             22      Gaps:      1

US-09-807-802A-15 (1-599) x AAF23750 (1-4679)
Qy      1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20
Db      2614 ACGGCTCCGGGAAAAAAGAGCGCGGTAGAGCACTCTCTGTGGAGCCAGACTCTCTCTCG 2673
Qy      21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40
Db      2674 GGAACCGGAAGCGCGGCCAGCAGCCTGCAAGAAAAAGATTGAATTTTGGTCAGACTGGA 2733
Qy      41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60
Db      2734 GAGCGAGACTCAGTACTGACCCCGCCAGCCTCTCGACAGCCACCAGCAGCCCCCTCTGGT 2793
Qy      61 ValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGly 80
Db      2794 CTGGGAACATAATACGATGGCTACAGGCAGTGGCGCACCATAATGGCAGACAATAACGAGGCG 2853
Qy      81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100
Db      2854 GCCGACGGAGTGGGTAATTCTCTCCGAAATTTGGCATTTGCCATTCCACATGGATGGGCGAC 2913
Qy      101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120
Db      2914 AGAGTCATCACCAACCAACCGACCCGAAACCTGGGCCCTGGCCACCTACAACACCACCTCTAC 2973
Qy      121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140
Db      2974 AAACAAAATTTCCAGCCCAATCA---GGAGCCTCGAACGACAATCACTACTTTGGCTACAGC 3030
Qy      141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160
```



CC acid insertion following the capsid amino acid at position 139, 161, 588  
CC or 657 in the VP1 capsid. The AAV vector comprises a capsid protein  
CC containing one or more amino acid insertions that ablate the ability of  
CC the vector to bind heparin-sulphate proteoglycan and allow the vector to  
CC use a cellular receptor not used by wild type AAV. Modified (I) are  
CC useful as vaccines to elicit immune responses to amino acids, where the  
CC response can be protective and/or therapeutic. (I) may be used to  
CC transfer a therapeutic peptide to a cancer cell, particularly to an  
CC ovarian cancer cell. The present sequence represents the adeno-associated  
CC virus 2 (AAV2) vector used in to make modified AAV2 vectors.

XX  
SQ Sequence 4679 BP; 1198 A; 1262 C; 1255 G; 964 T; 0 other;

Alignment Scores:

Pred. No.: 8.26e-219 Length: 4679  
Score: 2764.50 Matches: 496  
Percent Similarity: 89.98% Conservatives: 43  
Best Local Similarity: 82.80% Mismatches: 59  
Query Match: 85.04% Indels: 1  
DB: 24 Gaps: 1

US-09-807-802A-15 (1-599) x ABK89694 (1-4679)

QY 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSer 20  
Db 2614 ACGGCTCCGGGAAAAAAGAGCGCGGTAGAGACTCTCTGTGGAGCCAGACTCCTCTCG 2673  
QY 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40  
Db 2674 GGAACCGGAAAGCGCGGCCAGCAGCCTGCAAGAAAAAGATTGAATTTGGTCAGACTGGA 2733  
QY 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60  
Db 2734 GACGCAGACTCAGTACTGACCCCGCAGCCTCTCGGACAGCCACCGAGCCCTCTGTGT 2793  
QY 61 ValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGly 80  
Db 2794 CTGGGAACATAATACGATGGCTACAGGAGTGCGGCACCAATGGCAGACAATAACGAGGCG 2853  
QY 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100  
Db 2854 GCCGACGGAGTGGTAATCTCTCGGAAATTGGCATTCGATTCACATGGATGGCGGAC 2913  
QY 101 ArgValIleThrSerThrArgThrTrpAlaLeuProThrTrpAsnAsnHisLeuTyr 120  
Db 2914 AGAGTCATCACCCACGACACCGAACCTGGGCGCTGCCACCTACACCAACACCTCTAC 2973  
QY 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140  
Db 2974 AAACAAATTTCCAGCCAATCA--GGAGCCTCGAACGACAAATCACTACTTTGGCTACAGC 3030  
QY 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160  
Db 3031 ACCCCTTGGGGTATTTGACTTCAACAGATTCCACTGCCACTTTTTCACACGCTGACTGG 3090  
QY 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180  
Db 3091 CAAAGACTCATCAACAACAACAACTGGGATTCGACCCCAAGAGACTCAACTTCAAGCTCTTT 3150  
QY 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200  
Db 3151 AACATTCAAGTCAAGAGGTACGCGAGAATGACGGTACGACGAGATTGCCAATAACCTT 3210  
QY 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220  
Db 3211 ACCAGCAGGTTTCAAGTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTCCTCGGCTCG 3270  
QY 221 AlaHisGlnGlyCysLeuProProPheProAlaaspValPheMetIleProGlnTyrGly 240  
Db 3271 GCGCATCAAGGATGCTCCCGCGCTTCCAGCAGAC3TCTTTCATGGTGCCACAGTATGGA 3330  
QY 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260

Db 3331 TACCTCACCCCTGAACAACGGGAGTCAGGCAGTAGGACGCTCTTCAATTTTACTGCTGGAG 3390  
QY 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280  
Db 3391 TACTTTCTCTCAGATGCTGCTACCGGAAACAACATTACCTTCAGCTACACTTTTGAG 3450  
QY 281 GluValPropheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300  
Db 3451 GACGTTCTTTTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCT 3510  
QY 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320  
Db 3511 CTCATCGACCACTACCTGTATTACTTGAGCAGACAACAACACTCCAAGTGGAAACCACACG 3570  
QY 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340  
Db 3571 CAGTCAAGGCTTCAGTTTCTCAGGCCGCGGAGCGAGTGACATTCGGGACCAGTCTAGGAAC 3630  
QY 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360  
Db 3631 TGGCTTCTGGACCCCTGTACCGCCAGCAGCGAGTATCAAGACATCTGCGGATACAAC 3690  
QY 361 AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380  
Db 3691 AACAGTGAATACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTG 3750  
QY 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400  
Db 3751 GTGAATCCGGGCCCGCCATGGCAAGCCACAGGACGATGAAGAAAAGTTTTTCTCTCAG 3810  
QY 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420  
Db 3811 AGCGGGTTCATCTCTTGGGAAGCAAGGCTCAGAGAAAAACAATGTGGACATTGAAAAAG 3870  
QY 421 ValMetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440  
Db 3871 GTCATGATTACAGACGAGAGGAATCAGGACAACCAATCCCGTGGTACGGAGCAGTAT 3930  
QY 441 GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis 460  
Db 3931 GGTCTGTATCTACCAACCTCCAGAGAGGCAACAGACAGACGATACCGCAGATGTCAAC 3990  
QY 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 480  
Db 3991 ACACAAGCGGTTCTTCCAGGATGGTCTGGCAGGACAGAGATGTGTACTCTCAGGGGCC 4050  
QY 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500  
Db 4051 ATCTGGGCAAGATTCCACACAGCGGACGACATTTTACCCCTCTCCCTCATGGGTGA 4110  
QY 501 PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520  
Db 4111 TTCGACTTAAACACCCCTCCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAAT 4170  
QY 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540  
Db 4171 CCTTCGACCACCTTCAGTGGGCAAGTTTGTCTTCTTCATCACACAGTACTCCACGGGA 4230  
QY 541 GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro 560  
Db 4231 CAGGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAGGAAACAGCAACCGTGGAAATCCC 4290  
QY 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580  
Db 4291 GAAATTCAGTACACTTCCAACTACAACAGTCTGTATTATGTGGACTTTACTGTGGACACT 4350  
QY 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599  
Db 4351 AATGGCGTGATTTCAGAGCCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTG 4407

RESULT 12  
ABV76133  
ID ABV76133 standard; DNA; 4679 BP.





QY 421 ValMetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440  
|||||  
Db 3871 GTCATGATTACAGACGAAGAGGAATCAGACAACCAATCCCGTGGCTACGGAGCAGTAT 3930  
QY 441 GlyThrValAlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHis 460  
|||||  
Db 3931 GGTTCTGTATCTACCAACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAAC 3990  
QY 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTrpLeuGlnGlyPro 480  
Db 3991 ACACAAGGCGTCTTCCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCC 4050  
QY 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500  
Db 4051 ATCTGGGCAAGATTCCACACGCGACGACATTTTCACCCCTCTCCCTCATGGGTGGA 4110  
QY 501 PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520  
Db 4111 TTCGGACTTAACACCCCTCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAAT 4170  
QY 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540  
Db 4171 CCTTCGACCACCTTCAGTGGCGGAAAGTTTGTCTTCTTCATCACACAGTACTCCACGGGA 4230  
QY 541 GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro 560  
Db 4231 CAGGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAAAACAGCAACGCTGGAATCCC 4290  
QY 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580  
Db 4291 GAAATTCAGTACACTTCCAACCACTCAACAAGTCTGTTAATGTGGACTTACTGTGGACACT 4350  
QY 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599  
Db 4351 AATGGCGTGATTACAGAGCCTCGCCCAATTGGCACCAGATACCTGACTCGTAATCTG 4407

RESULT 13

ID AAH26326 standard; DNA; 7557 BP.  
XX  
AC AAH26326;  
XX  
DT 02-OCT-2001 (first entry)  
XX  
DE Adeno-associated virus packaging plasmid pfloxAAV.  
XX  
KW AAV; vector; pfloxAAV2; packaging cell line; gene therapy; ds.  
XX  
OS Adeno associated virus.  
XX  
PN WO20015361-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 26-JAN-2001; 2001WO-US02709.  
XX  
PR 26-JAN-2000; 2000US-0178536.  
XX  
PA (CHIR ) CHIRON CORP.  
XX  
PI Hardy SF;  
XX  
DR WPI; 2001-483239/52.  
XX  
PT Producing recombinant adeno-associated virus (rAAV) vector, by stably  
PT infecting eukaryotic host cell with rAAV vectors, helper adenovirus or  
PT herpes virus, recombinant adenovirus or herpes vectors -  
XX  
PS Example 1; Page 54-57; 63pp; English.  
XX  
CC The present sequence is that of adeno-associated virus (AAV)  
CC packaging plasmid pfloxAAV2. The following AAV packaging plasmids

CC were prepared: pKSrepap, a non-replicating control plasmid  
CC containing rep and cap genes but no inverted terminal repeats  
CC (ITRs); pAV2 (see AAH26325), containing a wild-type AAV genome; and  
CC pfloxAAV, a version of pAV2 with 2 loxP sites inserted such they  
CC flank the rep and cap genes and thus separate the ITRs from these  
CC genes. Each was combined with pCMV GFP (see AAH26324), a plasmid  
CC with a green fluorescent protein expressing recombinant AAV (rAAV)  
CC genome, and used to transfect 293 or Cre8 cells. After 6 hr, the  
CC medium was changed and wild-type adenovirus type 5 was added.  
CC Virus particles were harvested 3 days later. Plasmid pAV2  
CC produced mostly AAV and a low yield of GFP vector with no effect  
CC of Cre recombinase. pfloxAAV packaged as much GFP vector as the  
CC control plasmid and while the amount of GFP vector was not  
CC affected by Cre recombinase, the amount of floxAAV in the product  
CC was reduced to 1/10 of the 293 value by the action of Cre  
CC recombinase. In 293 cells it was evident that the GFP vector did  
CC not compete with wild-type AAV for replication or packaging, while  
CC the GFP vector was able to complete against floxAAV. Methods  
CC and compositions are provided for producing rAAV vector particles  
CC by: (a) introducing into a host cell (i) pfloxAAV, (ii) a  
CC recombinant viral vector encoding plasmid, and (iii) a plasmid  
CC encoding herpesvirus, cytomegalovirus or adenoviral functions, or  
CC a herpesvirus, cytomegalovirus or adenovirus itself, to produce  
CC flox AAV particles and rAAV particles; and (b) introducing into a  
CC second host cell (i) the rAAV particles or (a), (ii) a vector  
CC that directs expression of Cre, and (ii) a vector which directs  
CC expression of herpes virus, cytomegalovirus or adenovirus helper  
CC functions, such that rAAV vector particles are produced. The  
CC vectors are useful for in vivo or in vitro gene therapy and also  
CC for in vitro recombinant protein production.

XX  
SQ Sequence 7557 BP; 1970 A; 1940 C; 1953 G; 1694 T; 0 other;

Alignment Scores:

Pred. No.: 1.58e-218 Length: 7557  
Score: 2764.50 Matches: 496  
Percent Similarity: 89.98% Conservative: 43  
Best Local Similarity: 82.80% Mismatches: 59  
Query Match: 85.04% Indels: 1  
DB: 22 Gaps: 1

US-09-807-802A-15 (1-599) x AAH26326 (1-7557)

QY 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20  
Db 2642 ACGGCTCCGGGAAAAAAGAGCGGTAGAGCCTCTCTCTGTGGAGCCAGACTCTCTCTCG 2701  
QY 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40  
Db 2702 GGAACCGAAAGCGGGCCAGCAGCCTGCAAGAAAAGATTGAATTTGGTCAGACTGGA 2761  
QY 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60  
Db 2762 GACGCAGACTCAGTACTGACCCCGCAGCCTCTCGGACAGCCACGAGCCCTCTGGT 2821  
QY 61 ValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGly 80  
Db 2822 CTGGGAACCTAATACGATGGCTACAGGCAGTGGCGCAGTGGCCACCAATGGCAGACAATAACGAGGGC 2881  
QY 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100  
Db 2882 GCCGACGGAGTGGGTAATTCCTCGGGAATTTGGCATTTGCCGATTCACATGGATGGGCGAC 2941  
QY 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120  
Db 2942 AGAGTCATCACCACCAGCACCACCGAACCTGGGCCCTGGCCACCTTACAAACACCTCTAC 3001  
QY 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140  
Db 3002 AAACAAATTTCCAGCCCAATCA---GGAGCCTCGAACGACATCACTACTTTGGCTACAGC 3058  
QY 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160

Db 3059 ACCCCTTGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCCACCAGTGACTGG 3118

Qy 161 GlnArgLeuIleAsnAsnAenTirpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180

Db 3119 CAAGACTCATCAACAACAACACTGGGATTCCGACCCCAAGAGACTCAACITCAAGCTCTTT 3178

Qy 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200

Db 3179 AACATTCAAGTCAAGAGAGTCAAGCAGAATGACGGTACGACGACGATTGCCAATAACCTT 3238

Qy 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220

Db 3239 ACCAGCACGGTTTCAGGTGTTTACTGACTCGGAGTACCGACTCCCGTACGCTCGGCTCG 3298

Qy 221 AlaHisGlnGlyCysLeuProPropheProAlaAspValPheMetIleProGlnTyrGly 240

Db 3299 GCGCATCAAGGATGCTCCCGCGCTTCCAGCAGACGCTTTCATGGTGCCACAGTATGGA 3358

Qy 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260

Db 3359 TACCTCACCTGAACAACGGAGTCAAGGAGTACGGCAGTAGGACGCTCTTCATTTTACTGCTGGAG 3418

Qy 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280

Db 3419 TACTTTCCTTCTCAGATGTGGTACCGGAAACAACATTACCTTCAGCTACACTTTTGAG 3478

Qy 281 GluValPropheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300

Db 3479 GACGTTCTCTTCCACAGCAGTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCT 3538

Qy 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320

Db 3539 CTCATCGACCACTACCTGTATTACTTTGAGCAGAACAAACACTCCAAGTGGAAACCACACG 3598

Qy 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340

Db 3599 CAGTCAAGGCTTCAGTTTCTCAGGCCGGAGCGAGTGACATTCGGGACCAGTCTAGGAAC 3658

Qy 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360

Db 3659 TGGCTTCTCGACCCCTGTACCGCCAGCAGCGAGTATCAAGACATCTCGCGATAACAAC 3718

Qy 361 AsnSerAsnPheThrTirpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380

Db 3719 AACAGTGAATACTCGTGGAGTGGAGTACCAAGTACCACCTCAATGGCAGAGACTCTCTG 3778

Qy 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400

Db 3779 GTGAATCCGGCCCGCCATGGCAAGCCACAAAGGACGATGAAGAAAAGTTTTTCTCTCAG 3838

Qy 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420

Db 3839 AGCGGGGTTCTCATCTTTGGGAAGCAAGGCTCAGAGAAAACAATGTGGACATTGAAAAG 3898

Qy 421 ValMetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440

Db 3899 GTCATGATTACAGACGAAGAGGAAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTAT 3958

Qy 441 GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis 460

Db 3959 GGTCTGTATCTACCAACCTCCAGAGAGGCAACAGACAGCAGCTACCGCAGATGTCAAC 4018

Qy 461 AlaMetGlyAlaLeuProGlyMetValTirpGlnAspArgAspValTyrLeuGlnGlyPro 480

Db 4019 ACACAAGCGGTTCTTCCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCC 4078

Qy 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500

Db 4079 ATCTGGCAAGATTCCACACAGGACGGACATTTTTCACCCCTCTCCCTCATGGGTGGA 4138

Qy 501 PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520

Db 4139 TTCGGACTTAAACACCCCTCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGGAAT 4198

Qy 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540

Db 4199 CCTTCGACCACCTTCAGTCCGGCAAAGTTTGCTTCTTCATCACACAGTACTCCACGGGA 4258

Qy 541 GlnValSerValGluIleGluTirpGluLeuGlnLysGluAsnSerLysArgTirpAsnPro 560

Db 4259 CAGGTACGCGTGGAGATCGAGTGGAGCTGCAGAAGGAAACAGCAACGCTGGAATCCC 4318

Qy 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580

Db 4319 GAAATTCAGTACACTTCCAACTACAACAAGTCTGTTAATGTGGACTTACTGTGGACACT 4378

Qy 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599

Db 4379 AATGGCGTGTATTCAGAGCCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTG 4435

RESULT 14

AAH26325

ID AAH26325 standard; DNA; 8698 BP.

XX

AC AAH26325;

XX

DT 02-OCT-2001 (first entry)

XX

DE Wild-type adeno-associated virus genome plasmid pAV2.

XX

KW AAV; vector; pAV2; packaging cell line; gene therapy; ds.

XX

OS Adeno associated virus.

XX

PN WO200155361-A2.

XX

PD 02-AUG-2001.

XX

PF 26-JAN-2001; 2001WO-US02709.

XX

PR 26-JAN-2000; 2000US-0178536.

XX

PA (CHIR ) CHIRON CORP.

XX

PI Hardy SF;

XX

DR WPI; 2001-483239/52.

XX

PT Producing recombinant adeno-associated virus (rAAV) vector, by stably

PT infecting eukaryotic host cell with rAAV vectors, helper adenovirus or

PT herpes virus, recombinant adenovirus or herpes vectors -

XX

PS Example 1; Page 51-54; 63pp; English.

The present sequence is that of adeno-associated virus (AAV) packaging plasmid pAV2, which contains a wild-type AAV type 2 genome. The following AAV packaging plasmids were prepared: pXSrepCap, a non-replicating control plasmid containing rep and cap genes but no inverted terminal repeats (ITRs); pAV2; and pflxAAV (see AAH26326), a version of pAV2 in which 2 inserted loxP sites flank the rep and cap genes and thus separate the ITRs from these genes. Each was combined with pCMV GFP (see AAH26324), a plasmid with a green fluorescent protein expressing recombinant AAV (rAAV) genome, and used to transfect 293 or Cre8 cells. After 6 hr, the medium was changed and wild-type adenovirus type 5 was added. Virus particles were harvested 3 days later. Plasmid pAV2 produced mostly AAV and a low yield of GFP vector with no effect of Cre recombinase. pflxAAV packaged as much GFP vector as the control plasmid and while the amount of GFP vector was not affected by Cre recombinase, the amount of floxAAV in the product was reduced to 1/10 of the 293 value by the action of Cre recombinase. In 293 cells it was evident that the GFP vector did not compete with wild-type AAV for replication or packaging, while the GFP vector was able to compete against floxAAV. Methods and compositions are provided for producing rAAV vector particles by: (a) introducing into a host cell (i) pflxAAV, (ii) a



CC recombinant viral vector encoding plasmid, and (iii) a plasmid  
CC encoding herpesvirus, cytomegalovirus or adenoviral functions, or  
CC a herpesvirus, cytomegalovirus or adenovirus itself, to produce  
CC flox AAV particles and rAAV particles; and (b) introducing into a  
CC second host cell (i) the rAAV particles or (a), (ii) a vector  
CC that directs expression of Cre, and (ii) a vector which directs  
CC expression of herpes virus, cytomegalovirus or adenovirus helper  
CC functions, such that rAAV vector particles are produced. The  
CC vectors are useful for in vivo or in vitro gene therapy and also  
CC for in vitro recombinant protein production.  
XX  
SQ Sequence 8698 BP; 2117 A; 2368 C; 2306 G; 1907 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 1.91e-218 Length: 8698  
Score: 2764.50 Matches: 496  
Percent Similarity: 89.98% Conservative: 43  
Best Local Similarity: 82.80% Mismatches: 59  
Query Match: 85.04% Indels: 1  
DB: 22 Gaps: 1  
  
US-09-807-802A-15 (1-599) x AAH26325 (1-8698)  
  
QY 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20  
Db 2614 ACGGCTCCGGGAAAAAGAGCGGTAGAGCACTCTCTGTGGAGCCAGACTCTCTCTCG 2673  
  
QY 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40  
Db 2674 GGAACCGAAAGCGGGCCAGCAGCGCTGCAAGAAAAAGATTGAATTTGGTCAGACTGGA 2733  
  
QY 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60  
Db 2734 GACGCAGACTCAGTACCTGACCCCGAGCTCTCGGACAGCCAGCAGCCCTCTGCT 2793  
  
QY 61 ValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGly 80  
Db 2794 CTGGGAACTAATACGATGGCTACAGGCGAGTGGCGCACCAATGGCAGACAATAACGAGGSC 2853  
  
QY 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100  
Db 2854 GCCGACGGAGTGGGTAAATCTCTCGGGAAATGGCATTTGGATTCCACATGGATGGGCGAC 2913  
  
QY 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120  
Db 2914 AGAGTCATCACACCAGCAGCACCCGAACTTGGCCCTGCCCCACCTACAAACACCCTCTAC 2973  
  
QY 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140  
Db 2974 AAACAAATTTCCAGCCAATCA---GGAGCCTCGAACGACAAATCACTACTTTGGCTACAGC 3030  
  
QY 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160  
Db 3031 ACCCTTGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCACCACGTCGACTGG 3090  
  
QY 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180  
Db 3091 CAAAGACTCATCAACAACAACACTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTT 3150  
  
QY 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200  
Db 3151 AACATTCAAGTCAAGAGGTCAAGCAGAAATGACGGTACGACGACGATTGCCAATAACCTT 3210  
  
QY 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220  
Db 3211 ACCAGCAGGTTTCAAGTGTATTACTGACTCGAGTACCAGCTCCCGTACGTCCTCGGCTCG 3270  
  
QY 221 AlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGly 240  
Db 3271 GCGCATCAAGGATGCTCTCCCGCGTTCCAGCAGACGCTCTTCAATGGTGGCCACAGTATGGA 3330  
  
QY 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260  
Db 260

Db 3331 TACCTCACCTGAACAAACGGGAGTCAGGCAGTAGGACGCTCTTCAATTTTACTGCCTGGAG 3390  
QY 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280  
Db 3391 TACTTTCTCTCAGATGCTGCGTACCGGAAACAACCTTTACCTTCAGCTACACTTTTGAG 3450  
QY 281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300  
Db 3451 GACGTTCTTTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCT 3510  
QY 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320  
Db 3511 CTCATCGACCAGTACCTGTATTACTTGAGCAGACAACAACACTCCAAGTGGAAACCAACG 3570  
QY 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340  
Db 3571 CAGTCAAGGCTTCAGTTTCTCAGGCCGAGCGAGTGACATTCGGGACCAGTCTAGGAAC 3630  
QY 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360  
Db 3631 TGGCTTCTGGACCTGTATTACCGCCAGCAGCAGTATCAAAGACATCTCGGATAACAAC 3690  
QY 361 AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380  
Db 3691 AACAGTGAATACTCGTGAGCTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTG 3750  
QY 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400  
Db 3751 GTGAATCCGGGCCCGCCATGGCAAGCCACAGGACGATGAAGAAAGTTTTTCTCTCAG 3810  
QY 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420  
Db 3811 AGCGGGTTCTCATCTTTGGGAAGCAAGGCTCAGAGAAACAATGTGGACATTTGAAAG 3870  
QY 421 ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440  
Db 3871 GTCATGATTACAGACGAAGAGAAATCAGGACAAACAATCCCGTGGCTACGGAGCAGTAT 3930  
QY 441 GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis 460  
Db 3931 GGTTCGTATCTACCAACCTCCAGAGAGGCAACAGACAAGCAGTACCAGCAGATGTCAAC 3990  
QY 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 480  
Db 3991 ACACAAGGCGTTCTTCCAGGCGATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCC 4050  
QY 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500  
Db 4051 ATCTGGGCAAGATTCCACACACGCGGACATTTTCCACCCCTCTCCCTCATGGGTGA 4110  
QY 501 PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520  
Db 4111 TTGGGACTTAAACACCCCTCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAAT 4170  
QY 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540  
Db 4171 CCTTCGACCACCTTCAGTGGGCAAGTTTGTCTTCTCATCACACAGTACTCCACGGGA 4230  
QY 541 GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro 560  
Db 4231 CAGGTCAAGTGGAGATCGAGTGGGAGCTGCAGAAAGGAAACAGCAACGCTGGAATCCC 4290  
QY 561 GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn 580  
Db 4291 GAAATTCAGTACACTTCCAACACTACAACAAGTCTGTAAATGTGGACTTTTACTGTGGACT 4350  
QY 581 AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599  
Db 4351 AATGGCGTGTATTCAGAGCGCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTG 4407

RESULT 15  
AAT09008  
ID AAT09008 standard; DNA; 4680 BP.

```
XX AAT09008;
XX
DT 05-JUL-1996 (first entry)
XX
XX Wild-type adeno-associated virus 2 genome.
DE
XX
XX adeno-associated virus 2; AAV2; inverted terminal repeat; viral vector;
KW DNA delivery; immunodeficiency virus protein; immunity; human; simian;
KW neurological disorder; therapy; ss.
XX
OS Adeno-associated virus.
XX
XX WO9534670-A2.
PN
XX 21-DEC-1995.
PD
XX
XX 06-JUN-1995; 95WO-US07178.
PF
XX
XX 06-JUN-1994; 94US-0254358.
PR
XX
XX (CHIL-) CHILDRENS HOSPITAL INC.
PA
XX
XX Johnson PR;
PI
XX
XX WPI; 1996-049697/05.
DR
XX
XX Recombinant adeno-associated virus genome contg. protein encoding
PT DNA - flanked by inverted terminal repeats, for use in vaccines or
PT for treatment of neuro-degenerative disease
XX
XX Example 1; Page 27-29; 44pp; English.
PS
XX
XX The present sequence is that of the wild-type adeno-associated virus
CC (AAV) 2, a replication-deficient parvovirus. Cis-acting sequences
CC directing viral DNA replication (ori), encapsidation/packaging (pkg) and
CC host cell chromosome integration (int) are contained within the ITRs
CC (inverted terminal repeats of 145 nucleotides). When AAV infects a
CC human cell, the viral genome integrates into chromosome 19 resulting in
CC latent infection of the cell. Prodn. of infectious virus does not occur
CC unless the cell is infected with a helper virus (e.g., adenovirus or
CC herpes- virus). AAV possesses unique features that make it attractive as
CC a vector for delivering foreign DNA to cells. A vector including a
CC recombinant AAV genome contg. a simian immunodeficiency virus (SIV) rev
CC and envelop (gp160) gene cassette was constructed from an existing plasmid
CC designated psu201. The plasmid contains a modified wild-type AAV2
CC genome. Specifically, two XbaI sites were added via linker addition at
CC sequence positions 190 and 484. These sites are internal to 191 bp ITRs
CC which include the 145 bp ITRs of the AAV genome. The insertion of these
CC sites allows the complete removal of the internal 4.3 kb fragment contg.
CC the AAV rep-cap genes upon XbaI digestion of the plasmid. (See also
CC AAT09009-10).
XX
SQ Sequence 4680 BP; 1198 A; 1265 C; 1254 G; 963 T; 0 other;
```

```
Alignment Scores:
Pred. No.: 2.5e-215 Length: 4680
Score: 2722.50 Matches: 492
Percent Similarity: 89.00% Conservative: 42
Best Local Similarity: 82.00% Mismatches: 64
Query Match: 83.74% Indels: 2
DB: 17 Gaps: 1
```

US-09-807-802A-15 (1-599) x AAT09008 (1-4680)

```
QY 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20
DB 2614 ACGGTCGGGAAAGAGAGCGGTAGAGCACTCTCTCTGGAGCCAGACTCTCTCTCG 2673
QY 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40
DB 2674 GGAACCGGAAGCGGGCCAGACGCTGCAAGAAAGATTGAATTTTGGTCAGACTGGA 2733
```

```
QY 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60
DB 2734 GACGCAGACTCAGTACCTGACCCCGCAGCCTCTCGGACAGCCACCGAGCCCTCTGGT 2793
QY 61 ValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnGluGly 80
DB 2794 CTGGGAATAATACGATGGCTACAGCAGTGGCGCACCAATGGCAGACAATAACGAGGC 2853
QY 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100
DB 2854 GCCGACGAGTGGGTAAATTCCTCCGAAATGGCATTCGATTCCACATGGATGGCGAC 2913
QY 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTrpAsnAsnHisLeuTyr 120
DB 2914 AGAGTCATCACCCACAGCACCCGAACTGGGCCCTGCCACCTACAAACACCTCTAC 2973
QY 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140
DB 2974 AAACAAATTTCCAGCCAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGC 3030
QY 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTyr 160
DB 3031 ACCCCTTGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCCACACGTGACTGG 3090
QY 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180
DB 3091 CAAAGACTCATCAACAACACTGGGGATTCGACCCCAAGAGACTCAACTTCAAGCTCTTT 3150
QY 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200
DB 3151 AACATTCAAGTCAAAAGAGGTACCGCAGAAATGACGACGAGTTCGCAATTAACCTT 3210
QY 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220
DB 3211 ACCAGCACGGTTCAGGTGTTTACTGACTCGGAGTACCAGTCCCGTACGTCTCGGCTCG 3270
QY 221 AlaHisGlnGlyCysLeuProPropheProAlaAspValPheMetIleProGlnTyrGly 240
DB 3271 GCGCATCAAGGATGCTCCCGCGTTCCAGCAGACGCTCTTTCATGGTGCCACAGTATGGA 3330
QY 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260
DB 3331 TACCTCACCTGAACAACGGGAGTCAGGAGTAGGACGCTCTTCACTTACTGCTGGAG 3390
QY 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280
DB 3391 TACTTTCTCTCAGATGCTGCTACCGTAACCGAACAACCTTACCTTCAGCTACACTTTGAG 3450
QY 281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300
DB 3451 GACGTTCTCTTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCT 3510
QY 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320
DB 3511 CTCATCGACCACTACCTGTATTACTTGAGCAGACAACAACACTCCAAGTGGAAACCACCG 3570
QY 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340
DB 3571 CAGTCAAGGCTTCAGTCTTCTCAGGCCGAGCGAGTGACATTCGGACCACTAGTGAAC 3630
QY 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360
DB 3631 TGGCTTCTCGGACCTGTTACCGCCAGCAGCGAGTATCAAAGACATCTCGCGGATAACAAC 3690
QY 361 AsnSerAsnPheThrThrThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380
DB 3691 AACAGTGAATACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTG 3750
QY 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePhePromet 400
DB 3751 GTGAATCCGGGGCCCGCCATGGCAAGCCACAGGACGATGAAGAAAAGTTTTTCTCTCAG 3810
QY 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420
```

|    |      |  |      |
|----|------|--|------|
| Db | 3811 | AGCGGGTTCTCATCTTTGGGAAGCAAGGCTCAGAGAAAACAATGTGAACATTGAAAAG   | 3870 |
| Qy | 421  | ValMetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe | 440  |
| Db | 3871 | GTTCATGATTACAGACGAAGAGGAAATCGGAACAACCAATCCCGTGGCTACGGACAGTAT | 3930 |
| Qy | 441  | GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis | 460  |
| Db | 3931 | GGTTCTGTATCTACCAACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAAC | 3990 |
| Qy | 461  | AlaMetGlyAlaLeuProGlyMetValTyrGlnAspArgAspValTyrLeuGlnGlyPro | 480  |
| Db | 3991 | ACACAAGGCGTTCTTCCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCC  | 4050 |
| Qy | 481  | IleTyrAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly | 500  |
| Db | 4051 | ATCTGGGCAAGATTCCACACACGACGGACATTTTCACCCCTCTCCCTCATGGGTGA     | 4110 |
| Qy | 501  | PheGlyLeuLysAsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsn | 520  |
| Db | 4111 | TTCCGACCTTAAACACCCCTCTCCACAGATTCTCATCAAGAAACACCCCGGTACCTGC   | 4170 |
| Qy | 521  | ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly | 540  |
| Db | 4171 | CCTTCGACCACCTTCAGTGGGCAAGTTTGCTTCTTCATCACACAGTACTCCACGGGA    | 4230 |
| Qy | 541  | Gln-ValSerValGluIleGluTyrGluLeuGlnLysGluAsnSerLysArgTyrAsnPr | 560  |
| Db | 4231 | CACGGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAGGAGAAACAGCAACGCTGGAATCC  | 4290 |
| Qy | 560  | oGluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAs | 580  |
| Db | 4291 | CGAAATTCAGTACACTTCCAACTACAACAAGTCTGTAAATCGTGGACTTACCGTGGATAC | 4350 |
| Qy | 580  | nAsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu   | 599  |
| Db | 4351 | TAATGGCGTGTATTCCAGAGCTCGCCCCATTGGCCAGATACCTGACTCGTAATCTG     | 4408 |

Search completed: February 14, 2004, 17:22:02  
Job time : 533.871 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 14, 2004, 16:11:58 ; Search time 88.4559 Seconds  
(without alignments)  
2988.931 Million cell updates/sec

Title: US-09-807-802A-15  
Perfect score: 3251  
Sequence: 1 TAPGKRPEVQSQPEPDSS.....NNGLYTEPRPIGTRYLTRPL 599

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delopt 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09807802/runat\_11022004\_175609\_15955/app\_query.fasta\_1.2389  
-DB=Issued Patents NA -OPMT=fastap -SUFFIX=tni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09807802@cgn\_1\_131@runat\_11022004\_175609\_15955 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Match | Length | DB ID | Description       |
|------------|--------|-------|--------|-------|-------------------|
| 1          | 2801   | 86.2  | 8179   | 4     | US-09-438-268-5   |
| 2          | 2764.5 | 85.0  | 4072   | 4     | US-09-770-315-4   |
| 3          | 2764.5 | 85.0  | 7557   | 4     | US-09-770-315-3   |
| 4          | 2764.5 | 85.0  | 8698   | 4     | US-09-770-315-2   |
| 5          | 2722.5 | 83.7  | 4680   | 1     | US-08-254-358-1   |
| 6          | 2722.5 | 83.7  | 4680   | 1     | US-08-475-391-1   |
| 7          | 2722.5 | 83.7  | 4680   | 2     | US-08-709-609-1   |
| 8          | 2722.5 | 83.7  | 4680   | 5     | PCT-US95-07178-1  |
| 9          | 2398.5 | 73.8  | 8151   | 4     | US-09-438-268-2   |
| 10         | 1830.5 | 56.3  | 1800   | 4     | US-09-532-594B-17 |
| 11         | 1830.5 | 56.3  | 2208   | 4     | US-09-532-594B-5  |
| 12         | 1830.5 | 56.3  | 4767   | 4     | US-09-532-594B-1  |

|   |    |        |      |         |   |                      |                    |
|---|----|--------|------|---------|---|----------------------|--------------------|
| C | 13 | 1830.5 | 56.3 | 7214    | 4 | US-09-438-268-1      | Sequence 1, Appli  |
|   | 14 | 1690.5 | 52.0 | 1617    | 4 | US-09-532-594B-19    | Sequence 19, Appl  |
|   | 15 | 772    | 23.7 | 2271    | 4 | US-09-438-268-3      | Sequence 3, Appli  |
|   | 16 | 247    | 7.6  | 1740    | 1 | US-07-969-213-1      | Sequence 1, Appli  |
|   | 17 | 231.5  | 7.1  | 2062    | 6 | 5223424-3            | Patent No. 5223424 |
|   | 18 | 220    | 6.8  | 1820    | 6 | 5223424-12           | Patent No. 5223424 |
|   | 19 | 209    | 6.4  | 5049    | 1 | US-08-336-345-1      | Sequence 1, Appli  |
|   | 20 | 209    | 6.4  | 5049    | 1 | US-08-336-345-2      | Sequence 2, Appli  |
|   | 21 | 209    | 6.4  | 5049    | 2 | US-08-647-655-1      | Sequence 1, Appli  |
|   | 22 | 209    | 6.4  | 5049    | 2 | US-08-647-655-2      | Sequence 2, Appli  |
|   | 23 | 203    | 6.2  | 2254    | 3 | US-08-552-369-1      | Sequence 1, Appli  |
|   | 24 | 192    | 5.9  | 1752    | 3 | US-09-022-949-1      | Sequence 1, Appli  |
|   | 25 | 184    | 5.7  | 1755    | 2 | US-08-317-785-1      | Sequence 1, Appli  |
| C | 26 | 131    | 4.0  | 4188    | 4 | US-09-252-991A-13774 | Sequence 13774, A  |
|   | 27 | 131    | 4.0  | 8211    | 4 | US-09-252-991A-13656 | Sequence 13656, A  |
| C | 28 | 123.5  | 3.8  | 1896    | 1 | US-08-198-446B-8     | Sequence 8, Appli  |
| C | 29 | 123.5  | 3.8  | 1896    | 2 | US-08-870-693-8      | Sequence 8, Appli  |
| C | 30 | 120    | 3.7  | 4403765 | 3 | US-09-103-840A-2     | Sequence 2, Appli  |
| C | 31 | 119.5  | 3.7  | 4411529 | 3 | US-09-103-840A-1     | Sequence 1, Appli  |
|   | 32 | 119    | 3.7  | 3288    | 4 | US-09-107-532A-201   | Sequence 201, App  |
| C | 33 | 119    | 3.7  | 4214    | 4 | US-09-122-135-1      | Sequence 1, Appli  |
|   | 34 | 117.5  | 3.6  | 3906    | 4 | US-09-996-243-83     | Sequence 83, Appl  |
|   | 35 | 114    | 3.5  | 3155    | 2 | US-08-231-193A-13    | Sequence 13, Appl  |
|   | 36 | 114    | 3.5  | 3155    | 2 | US-08-486-273A-13    | Sequence 13, Appl  |
|   | 37 | 114    | 3.5  | 3155    | 3 | US-08-480-474-13     | Sequence 13, Appl  |
|   | 38 | 114    | 3.5  | 3155    | 3 | US-08-940-086A-13    | Sequence 13, Appl  |
|   | 39 | 114    | 3.5  | 3155    | 4 | US-08-940-035A-13    | Sequence 13, Appl  |
|   | 40 | 114    | 3.5  | 3155    | 4 | US-08-935-105A-13    | Sequence 13, Appl  |
|   | 41 | 114    | 3.5  | 3155    | 4 | US-09-648-797-13     | Sequence 13, Appl  |
|   | 42 | 114    | 3.5  | 3155    | 4 | US-09-386-123-13     | Sequence 13, Appl  |
|   | 43 | 114    | 3.5  | 3211    | 2 | US-08-231-193A-21    | Sequence 21, Appl  |
|   | 44 | 114    | 3.5  | 3211    | 2 | US-08-486-273A-21    | Sequence 21, Appl  |
|   | 45 | 114    | 3.5  | 3211    | 3 | US-08-480-474-21     | Sequence 21, Appl  |

ALIGNMENTS

RESULT 1  
US-09-438-268-5  
; Sequence 5, Application US/09438268  
; Patent No. 6491907  
; GENERAL INFORMATION:  
; APPLICANT: Rabinowitz, Joseph E.  
; APPLICANT: Samulski, Richard J  
; APPLICANT: Xiao, Weidong  
; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING  
; TITLE OF INVENTION: THE SAME  
; FILE REFERENCE: 5470-186  
; CURRENT APPLICATION NUMBER: US/09/438,268  
; CURRENT FILING DATE: 1999-11-10  
; EARLIER APPLICATION NUMBER: 60/107,840  
; EARLIER FILING DATE: 1998-11-10  
; EARLIER APPLICATION NUMBER: 60/123,651  
; EARLIER FILING DATE: 1999-03-10  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 8179  
; TYPE: DNA  
; ORGANISM: Virus  
US-09-438-268-5

Alignment Scores:  
Pred. No.: 8.59e-261 Length: 8179  
Score: 2801.00 Matches: 509  
Percent Similarity: 90.33% Conservative: 33  
Best Local Similarity: 84.83% Mismatches: 56  
Query Match: 86.16% Indels: 2  
DB: 4 Gaps: 2

US-09-807-802A-15 (1-599) x US-09-438-268-5 (1-8179)

QY 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSer 20

```

Db      2544  |||||ACGGCTCCGGGAAAAAGAGCGGTAGAGCACTCTCCTGTGGAGCCAGACTCCTCTCTCG 2603
Qy      21    GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40
Db      2604  GGAACCGAAAGCGCGCCAGCAGCTGCAAGAAAGAAAGATTGAATTTTGGTCAGACTGGA 2663
Qy      41    AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60
Db      2664  GACGCAGACTCAGTACCTGACCCCCAGCCTCTCGGACAGCCACGAGCCCCCTCTGCT 2723
Qy      61    ValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGly 80
Db      2724  CTGGGAACATAATACGATGGCTACAGGCAGTGGCGCACCAATGCGAGACAATAACGAGGCG 2783
Qy      81    AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100
Db      2784  GCCGACGAGTGGGTAAATTCCTCCGGAAATTGGCATTTGCCATTCCACATGGATGGGCGAC 2843
Qy      101   ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120
Db      2844  AGAGTCATCACCCACAGCACCCGAACTGGGCCCTGCCACCTTACAAACAACCCCTCTAC 2903
Qy      121   LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140
Db      2904  AAACAAATTTCCAGCCAATCA---GGAGCCTCGAACGACAAATCACTACTTTGGCTACAGC 2960
Qy      141   ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160
Db      2961  ACCCCTTGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCACCACCGTACTGG 3020
Qy      161   GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180
Db      3021  CAAAGACTCATCAACAACAACCTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTT 3080
Qy      181   AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200
Db      3081  AACATTCAAGTCAAGAGGTCAAGCAGATGACGGTACGACGACGATGGCCAATAACCTT 3140
Qy      201   ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220
Db      3141  ACCAGCACGGTTTCAAGTGTCTTACTGACTCGAGTACCAGTCCCGTACGTGCTCGGGTCG 3200
Qy      221   AlaHisGlnGlyCysLeuProPropheProAlaAspValPheMetIleProGlnTyrGly 240
Db      3201  GCGCACCAAGGCTGTCTCCGCCGTTTCCAGCGGACGTCTTTCATGGTCCCTCAGTATGGA 3260
Qy      241   TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260
Db      3261  TACCTCACCCCTGAACAACGGAAGTCAAGCGGTGGGACGCTCATCTCTTTTACTGCTGGAG 3320
Qy      261   TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280
Db      3321  TACTTCCCTTCGCAGATGCTAAGGACTGGAATAAATCTTCCAAATTCAGCTATACCTTCGAG 3380
Qy      281   GluValPropheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300
Db      3381  GATGTACCTTTTCACAGCAGCTACGCTCACAGCCAGAGTTTGGATCGCTTGATGAATCCT 3440
Qy      301   LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGln---AsnGlnSerGlySerAla 319
Db      3441  CTTATTGATCAGTATCTGTACTACTCTGAACAGAACGCAAGGAACAACCTCTGGAACAACC 3500
Qy      320   GlnAsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLys 339
Db      3501  AACCAATCACGGCTGCTTTTAGCCAGGCTGGGCTCAGTCTATGTCTTTCAGGCGCCAGA 3560
Qy      340   AsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsn 359
Db      3561  AATTGGCTACCTTGGGCCCTGCTACCGGCAACAGAGACTTTCAAAGACTGCTAACGACAAC 3620
Qy      360   AsnAsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSer 379

```

```

Db      3621  AACAAACAGTAACATTTCCTTGGACAGCGGCCAGCAAAATATCATCTCAATGGCCGCGACTCG 3680
Qy      380   IleIleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePhePro 399
Db      3681  CTGGTGAATCCAGGACCAGCTATGGCCAGTCACAAGGACGATGAAGAAAAATTTTCCCT 3740
Qy      400   MetSerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAsp 419
Db      3741  ATGCACGGCAATCTAATATTTTGGCAAAGAGGGACAACGGCAAGTAACGCAGAAATTAGAT 3800
Qy      420   AsnValMetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArg 439
Db      3801  AATGTAATGATTACGGATGAAGAAGAGATTTCGTACCACCAATCCTGTGGCAACACAGACAG 3860
Qy      440   PheGlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspVal 459
Db      3861  TATGGAACCTGTGGCAATAAATCTTGACAGAGCTCAAAATACAGCTCCCACGACTGGAACTGTC 3920
Qy      460   HisAlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGly 479
Db      3921  AATCATCAGGGGCCCTTACCTGGCATGGTGTGGCAAGATCGTGACGTGTACCTTCAAGGA 3980
Qy      480   ProIleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGly 499
Db      3981  CCTATCTGGCAAAAGATTCTCACACGGATGGACACTTTCATCTCTCTCTGTATGGGA 4040
Qy      500   GlyPheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAla 519
Db      4041  GGCTTTGGACTGAACAACATCCGCCTCTCAAATCATGATCAAAAATACTCCGGTACCTGCG 4100
Qy      520   AsnProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThr 539
Db      4101  AATCCTTCGACCACTTCAGTGGCGCAAGTTTGTCTTCTCATCACACAGTACTCCACG 4160
Qy      540   GlyGlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsn 559
Db      4161  GGACAGGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAGGAAACAGCAACCGCTGGAAT 4220
Qy      560   ProGluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAsp 579
Db      4221  CCGAAATTCAGTACACTTCCAACCTACAACAGTCTGTAAATCGTGGACTTACCGTGGAT 4280
Qy      580   AsnAsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599
Db      4281  ACTAATGGCGTGTATTCAGAGCCTCGCCCAATTGGCACCAAGATACCTGACTCGTAATCTG 4340

```

```

RESULT 2
US-09-770-315-4
; Sequence 4, Application US/09770315
; Patent No. 6429001
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; TITLE OF INVENTION: Recombinant AAV Packaging Systems
; FILE REFERENCE: 20263-501
; CURRENT APPLICATION NUMBER: US/09/770,315
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,536
; PRIOR FILING DATE: 2000-01-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 4072
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: recombinant DNA
US-09-770-315-4

Alignment Scores:
Pred. No.:          9.78e-258          Length:          4072
Score:              2764.50           Matches:          496
Percent Similarity: 89.98%            Conservative:     43
Best Local Similarity: 82.80%          Mismatches:       59

```

|   |        |   |      |
|---|--------|---|------|
| Query Match:  | 85.04% | Indels:   | 1    |
| DB:   | 4      | Gaps:   | 1    |
| US-09-807-802A-15 (1-599) x US-09-770-315-4 (1-4072)    |        |   |      |
| QY  | 1      | ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer  | 20   |
| Db  | 1895   | ACGGCTCCGGGAAAAAGAGGCGGTAGAGCACTCTCCTGTGGAGCCAGACTCCTCCTCG    | 1954 |
| QY  | 21     | GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly  | 40   |
| Db  | 1955   | GGAACCGGAAGCGGGCCAGCAGCGCTGCAAGAAAAAGATTGAATTTTGGTCAGACTGGA   | 2014 |
| QY  | 41     | AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla  | 60   |
| Db  | 2015   | GACGCAGACTCAGTACCTGACCCCGAGCCTCTCGGACAGCCACGAGCCCCCTCTGGT     | 2074 |
| QY  | 61     | ValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGly     | 80   |
| Db  | 2075   | CTGGGAACCTAATACGATGGCTACAGGCAGTGGCGCACCAATGGCAGACAAATAACGAGGC | 2134 |
| QY  | 81     | AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp  | 100  |
| Db  | 2135   | GCCGCGGAGTGGGTAAATTCCTCGGAAATTTGGCATTTGCGATTCCCATGGATGGSGGAC  | 2194 |
| QY  | 101    | ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr  | 120  |
| Db  | 2195   | AGAGTCATCACACCAGCACCCGGAACCTGGGCCCTGCCACCTACAACAACCACTCTAC    | 2254 |
| QY  | 121    | LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer  | 140  |
| Db  | 2255   | AAACAAATTTCCAGCCAATCA---GGAGCCTCGAACGACAAATCACTATTTTGGCTACAGC | 2311 |
| QY  | 141    | ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp  | 160  |
| Db  | 2312   | ACCCCTTGGGGGTATTTTGTACTTCAACAGATTCCACTGCCACTTTTCAACCAGTGACTGG | 2371 |
| QY  | 161    | GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe  | 180  |
| Db  | 2372   | CAAGACTCATCAACAACAACACTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTT | 2431 |
| QY  | 181    | AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu  | 200  |
| Db  | 2432   | AACATTCAAGTCAAAGAGGTCAAGCAGAAATGACGGTACGACGACGATTGCCAATAACCTT | 2491 |
| QY  | 201    | ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer  | 220  |
| Db  | 2492   | ACCAGCACGGTTACGGTGTTTACTGACTCGGAGTACCAGTCCCGTACCGTCCCTCGGCTCG | 2551 |
| QY  | 221    | AlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGly  | 240  |
| Db  | 2552   | GCGCATCAAGGATGCCCTCCCGCGTTCCAGCAGACGTCCTTCATGGTGCCACAGTATGGA  | 2611 |
| QY  | 241    | TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu  | 260  |
| Db  | 2612   | TACCTCACCTTGAACAACGGGAGTCAGGAGTAGGACGCTCTTCATTTTACTGCTGGAG    | 2671 |
| QY  | 261    | TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu  | 280  |
| Db  | 2672   | TACTTTCTCTCAGATGCTGCGTACCGGAAACAACCTTTACCTTCAGCTACACTTTTGAG   | 2731 |
| QY  | 281    | GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro  | 300  |
| Db  | 2732   | GACGTTCTCTTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCT  | 2791 |
| QY  | 301    | LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln  | 320  |
| Db  | 2792   | CTCATCGACCAGTACCTGTATTACTTGAGCAGAACAAACACTCCAAGTGGAAACCAACAG  | 2851 |
| QY  | 321    | AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn  | 340  |
| Db  | 2852   | CAGTCAAGGCTTCAGTTTCTCAGGCCGGAGCGAGTGACATTCGGGACCAAGTCTAGGAAC  | 2911 |
| QY  | 341    | TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn  | 360  |
| Db  | 2912   | TGGCTTCTTGGACCCTGTTTACCGCCAGCAGCGAGTATCAAAGACATCTGCGGATAACAAC | 2971 |
| QY  | 361    | AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle  | 380  |
| Db  | 2972   | AACAGTGAATACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTG  | 3031 |
| QY  | 381    | IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet  | 400  |
| Db  | 3032   | GTGAATCCGGGCCCCGCCCATGGCAAGCCACAAGGACGATGAAGAAAAGTTTTTTCCTCAG | 3091 |
| QY  | 401    | SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn  | 420  |
| Db  | 3092   | AGCGGGTTCTCATCTTTGGGAAGCAAGGCTCAGAGAAAAACAATGTGGACATTGAAAAAG  | 3151 |
| QY  | 421    | ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe     | 440  |
| Db  | 3152   | GTCAATGATTACAGACGAAGAGGAATCAGGACAAACCAATCCCGTGGCTACGGAGCAGTAT | 3211 |
| QY  | 441    | GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis  | 460  |
| Db  | 3212   | GGTTCTGTATCTACCAACCTCCAGAGAGGCAACAGACAGCTACCCGACAGATGTCAAC    | 3271 |
| QY  | 461    | AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro  | 480  |
| Db  | 3272   | ACACAAGGCGTTCTTCCAGGCATGGTCTGCGCAGCAGACAGATGTGTACCTTCAGGGGCC  | 3331 |
| QY  | 481    | IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly  | 500  |
| Db  | 3332   | ATCTGGGCAAGATTCCACACACGCGGACATTTTCACCCCTCTCCCTCATGGGTGGA      | 3391 |
| QY  | 501    | PheGlyLeuLysAsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsn  | 520  |
| Db  | 3392   | TTCGGACTTAAACACCCCTCTCCACAGATTCTCATCAAGAACACCCCGGTACTTGGGAAT  | 3451 |
| QY  | 521    | ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly  | 540  |
| Db  | 3452   | CCTTCGACCCCTTCAGTGGGCAAGTTTGTCTTCTTCATCACACAGTACTCCACGGGA     | 3511 |
| QY  | 541    | GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro  | 560  |
| Db  | 3512   | CAGGTCAAGCTGGAGATCGAGTGGGAGCTGCAGAAAGGAAAAACAGCAACGCTGGAATCCC | 3571 |
| QY  | 561    | GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn  | 580  |
| Db  | 3572   | GAATTCAGTACACTTCCAACACTACAACAAGTCTGTTAATGTGGACTTTACTGTGGACACT | 3631 |
| QY  | 581    | AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu     | 599  |
| Db  | 3632   | AATGGCGTGATTTCAGAGCCTCGCCCCCATTTGGCACCAGATACCTGACTCGTAATCTG   | 3688 |
| RESULT 3  |        |   |      |
| US-09-770-315-3   |        |   |      |
| ; Sequence 3, Application US/09770315                   |        |   |      |
| ; Patent No. 6429001                                    |        |   |      |
| ; GENERAL INFORMATION:                                  |        |   |      |
| ; APPLICANT: Chiron Corporation                         |        |   |      |
| ; TITLE OF INVENTION: Recombinant AAV Packaging Systems |        |   |      |
| ; FILE REFERENCE: 20263-501                             |        |   |      |
| ; CURRENT APPLICATION NUMBER: US/09/770,315             |        |   |      |
| ; CURRENT FILING DATE: 2001-01-26                       |        |   |      |
| ; PRIOR APPLICATION NUMBER: US 60/178,536               |        |   |      |
| ; PRIOR FILING DATE: 2000-01-26                         |        |   |      |
| ; NUMBER OF SEQ ID NOS: 8                               |        |   |      |
| ; SOFTWARE: FastSeq for Windows Version 3.0             |        |   |      |
| ; SEQ ID NO 3   |        |   |      |
| ; LENGTH: 7557  |        |   |      |
| ; TYPE: DNA   |        |   |      |
| ; ORGANISM: Unknown                                     |        |   |      |
| ; FEATURE:  |        |   |      |
| ; OTHER INFORMATION: recombinant DNA                    |        |   |      |
| US-09-770-315-3   |        |   |      |



Alignment Scores:

Pred. No.: 2,62e-257 Length: 7557  
Score: 2764.50 Matches: 496  
Percent Similarity: 89.98% Conservative: 43  
Best Local Similarity: 82.80% Mismatches: 59  
Query Match: 85.04% Indels: 1  
DB: 4 Gaps: 1

US-09-807-802A-15 (1-599) x US-09-770-315-3 (1-7557)

|    |      |  |      |
|----|------|--|------|
| QY | 1    | ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSer        | 20   |
| Db | 2642 | ACGGCTCCGGGAAAAAAGAGCGCGGTAGAGCACTCTCCTGTGGAGCCAGACTCCTCTCG      | 2701 |
| QY | 21   | GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly     | 40   |
| Db | 2702 | GGAACCGAAAGCGGCCAGCAGCCTGCAAGAAAAAGATTGAATTTTGGTCAGACTGGA        | 2761 |
| QY | 41   | AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla     | 60   |
| Db | 2762 | GAGCGAGACTCAGTACTGACCCCGAGCCTCTCGACAGCCACCAGCAGCCCTCTGGT         | 2821 |
| QY | 61   | ValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnGluGly        | 80   |
| Db | 2822 | CTGGGAACATAATACGATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGCG    | 2881 |
| QY | 81   | AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp     | 100  |
| Db | 2882 | GCCGACGGAGTGGGTAATCTCTCGGAAATTGGCATTCGATTCCACATGGATGGCGAC        | 2941 |
| QY | 101  | ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTrpAsnAsnHisLeuTyr     | 120  |
| Db | 2942 | AGAGTCATCACCCAGCACCCGAACTCTGGGCCCTGCCACCTACAACAACACCTCTAC        | 3001 |
| QY | 121  | LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer     | 140  |
| Db | 3002 | AAACAAATTTCCAGCCCAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGC     | 3058 |
| QY | 141  | ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp     | 160  |
| Db | 3059 | ACCCCTTGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCCACCGTGACTGG       | 3118 |
| QY | 161  | GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe     | 180  |
| Db | 3119 | CAAGACTCATCAACAACAACCTGGGATTCGACCCCAAGAGACTCAACTTCAAGCTCTTT      | 3178 |
| QY | 181  | AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu     | 200  |
| Db | 3179 | AACATTCAAGTCAAGAGTCAAGCAGAAATGACGTTACGAGTACGACGAGATTGCCAATAACCTT | 3238 |
| QY | 201  | ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer     | 220  |
| Db | 3239 | ACCAGCAGGTTTCAGGTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTCCTCGGCTCG     | 3298 |
| QY | 221  | AlaHisGlnGlyCysLeuProPropheProAlaAspValPheMetIleProGlnTyrGly     | 240  |
| Db | 3299 | GCGCATCAAGGATGCCTCCGCGCTTCCCAGCAGACGCTCTTCATGGTGCCACAGTATGGA     | 3358 |
| QY | 241  | TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu     | 260  |
| Db | 3359 | TACCTCACCTTGAACAACGGGAGTCAGGCAGTAGGACGCTCTTCAATTTACTGCGCTGGAG    | 3418 |
| QY | 261  | TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu     | 280  |
| Db | 3419 | TACTTCTCTCAGATGCTGCGTACCGGAAACAACCTTTACCTTCAGCTACACTTTTGAG       | 3478 |
| QY | 281  | GluValPropheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro     | 300  |
| Db | 3479 | GAGGTTCTTTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCT      | 3538 |
| QY | 301  | LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln     | 320  |

|    |      |   |      |
|----|------|---|------|
| Db | 3539 | CTCATCGACCAGTACCTGTATTACTTGAGCAGAAACAACACTCCAAGTGGAAACCACCACG | 3598 |
| QY | 321  | AsnLysAspLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn     | 340  |
| Db | 3599 | CAGTCAAGGCTTCAGTTTTTCAGGCCGAGCGAGTGACATTCGGGACCAGTCTAGGAAC    | 3658 |
| QY | 341  | TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn  | 360  |
| Db | 3659 | TGGCTTCTGGACCCCTGTTACCGCCAGCAGCGAGTATCAAGACATCTGCGGATAACAAC   | 3718 |
| QY | 361  | AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle  | 380  |
| Db | 3719 | AACAGTGAATACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTG  | 3778 |
| QY | 381  | IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePhePromet  | 400  |
| Db | 3779 | GTGAATCCGGGCCCGCCATGGCAAGCCACAAGGACGATGAAGAAAAAGTTTTTTCCTCAG  | 3838 |
| QY | 401  | SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn  | 420  |
| Db | 3839 | AGCGGGTTCATCTTTGGGAAGCAAGGCTCAGAGAAAAACAATGTGGACATTGAAAAAG    | 3898 |
| QY | 421  | ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe     | 440  |
| Db | 3899 | GTCTGATTACAGACGAAGAGGAAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTAT   | 3958 |
| QY | 441  | GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis  | 460  |
| Db | 3959 | GGTTCGTATCTACCAACCTCCAGAGAGGCAACAGACAAGCAGTACCGCAGATGTCAAC    | 4018 |
| QY | 461  | AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro  | 480  |
| Db | 4019 | ACACAAGGCGTCTTCCAGGCAATGCTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCC   | 4078 |
| QY | 481  | IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly  | 500  |
| Db | 4079 | ATCTGGCAAAAGATTCCACACACGACGACGACATTTTCAACCCCTCTCCCTCATGGTGA   | 4138 |
| QY | 501  | PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn     | 520  |
| Db | 4139 | TTCGGACTTAAACACCTCTCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAAT  | 4198 |
| QY | 521  | ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly  | 540  |
| Db | 4199 | CCTTCGACCACCTTCAGTGGCGCAAGTTTGCTTCTTCATCACACAGTACTCCACGGGA    | 4258 |
| QY | 541  | GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro  | 560  |
| Db | 4259 | CAGGTACGGTGGAGATCGAGTGGAGCTGCAGAGGAAAAACAGCAACGCTGGAATCCC     | 4318 |
| QY | 561  | GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn  | 580  |
| Db | 4319 | GAATTCAGTACACTTCCAACACTACAACAAGTCTGTTAATGTGGACTTTACTGTGGACACT | 4378 |
| QY | 581  | AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu     | 599  |
| Db | 4379 | AATGGCGTGTATTTCAGAGCCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTG    | 4435 |

RESULT 4  
US-09-770-315-2  
; Sequence 2, Application US/09770315  
; Patent No. 6429001  
; GENERAL INFORMATION:  
; APPLICANT: Chiron Corporation  
; TITLE OF INVENTION: Recombinant AAV Packaging Systems  
; FILE REFERENCE: 20263-501  
; CURRENT APPLICATION NUMBER: US/09/770,315  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: US 60/178,536  
; PRIOR FILING DATE: 2000-01-26  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 2

|  |      |  |      |               |  |      |  |  |  |
|--|------|--|------|---------------|--|------|--|--|--|
| ; LENGTH: 8698                                       |      |  |      |               |  |      |  |  |  |
| ; TYPE: DNA  |      |  |      |               |  |      |  |  |  |
| ; ORGANISM: Unknown                                  |      |  |      |               |  |      |  |  |  |
| ; FEATURE:   |      |  |      |               |  |      |  |  |  |
| ; OTHER INFORMATION: recombinant DNA                 |      |  |      |               |  |      |  |  |  |
| US-09-770-315-2                                      |      |  |      |               |  |      |  |  |  |
| Alignment Scores:                                    |      |  |      |               |  |      |  |  |  |
| Pred. No.:   |      | 3.28e-257  |      | Length:       |  | 8698 |  |  |  |
| Score:   |      | 2764.50  |      | Matches:      |  | 496  |  |  |  |
| Percent Similarity:                                  |      | 89.98%   |      | Conservative: |  | 43   |  |  |  |
| Best Local Similarity:                               |      | 82.80%   |      | Mismatches:   |  | 59   |  |  |  |
| Query Match:   |      | 85.04%   |      | Indels:       |  | 1    |  |  |  |
| DB:  |      | 4  |      | Gaps:         |  | 1    |  |  |  |
| US-09-807-802A-15 (1-599) x US-09-770-315-2 (1-8698) |      |  |      |               |  |      |  |  |  |
| QY   | 1    | ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer   | 20   |               |  |      |  |  |  |
| Db   | 2614 | ACGGCTCCGGGAAAAAGAGCGGTAGAGCACTCTCTGTGGAGCCAGACTCTCTCTCG       | 2673 |               |  |      |  |  |  |
| QY   | 21   | GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly   | 40   |               |  |      |  |  |  |
| Db   | 2674 | GGAACCGGAAGCGGGCCAGCAGCCTGCAAGAAAAAGATTGAATTTGGTCAGACTGGA      | 2733 |               |  |      |  |  |  |
| QY   | 41   | AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla   | 60   |               |  |      |  |  |  |
| Db   | 2734 | GACGCAGACTCAGTACCTGACCCCGGCTCTCGGACAGCCACCAGCAGCCCTCTGTT       | 2793 |               |  |      |  |  |  |
| QY   | 61   | ValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGly      | 80   |               |  |      |  |  |  |
| Db   | 2794 | CTGGGAACCTAATACGATGGCTACAGGCACTGGCGCAACCAATGGCAGACAATAACGAGGGC | 2853 |               |  |      |  |  |  |
| QY   | 81   | AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp   | 100  |               |  |      |  |  |  |
| Db   | 2854 | GCCGACGGAGTGGGTAATTCCTCGGGAATTTGGCAATTCGATTCACATGGATGGCGAC     | 2913 |               |  |      |  |  |  |
| QY   | 101  | ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr   | 120  |               |  |      |  |  |  |
| Db   | 2914 | AGAGTCATCACCACCGACCCGAACTTGGGCCCTGCCACCTACAACCAACCACTCTAC      | 2973 |               |  |      |  |  |  |
| QY   | 121  | LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer   | 140  |               |  |      |  |  |  |
| Db   | 2974 | AAACAAATTTCCAGCCCAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGC   | 3030 |               |  |      |  |  |  |
| QY   | 141  | ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp   | 160  |               |  |      |  |  |  |
| Db   | 3031 | ACCCCTTGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTACCACCGTGACTGG    | 3090 |               |  |      |  |  |  |
| QY   | 161  | GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe   | 180  |               |  |      |  |  |  |
| Db   | 3091 | CAAAAGACTCATCAACAACAATGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTT  | 3150 |               |  |      |  |  |  |
| QY   | 181  | AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu   | 200  |               |  |      |  |  |  |
| Db   | 3151 | AACATTCAAGTCAAAGAGTCAACGAGAATGACGGTACGACGACGATTGCCAATAACCTT    | 3210 |               |  |      |  |  |  |
| QY   | 201  | ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer   | 220  |               |  |      |  |  |  |
| Db   | 3211 | ACCAGCAGGTTACGGTGTCTTACTGACTCGGAGTACCACTCCCGTACGTCCTCGGCTCG    | 3270 |               |  |      |  |  |  |
| QY   | 221  | AlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGly   | 240  |               |  |      |  |  |  |
| Db   | 3271 | GCGCATCAAGGATGCCTCCGCGGTTCCAGCAGACGCTCTTCATGGTGCCACAGTATGGA    | 3330 |               |  |      |  |  |  |
| QY   | 241  | TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu   | 260  |               |  |      |  |  |  |
| Db   | 3331 | TACCTCACCTGAACAACGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCCTGGAG     | 3390 |               |  |      |  |  |  |
| QY   | 261  | TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu   | 280  |               |  |      |  |  |  |
| Db   | 3391 | TACTTTCCTTCTCAGATGCTGCGTACCGGAAACAACITTTACCTTCAGCTACACTTTTGAG  | 3450 |               |  |      |  |  |  |

|    |      |  |      |
|----|------|--|------|
| QY | 281  | GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro   | 300  |
| Db | 3451 | GACGTTCTTTTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCT   | 3510 |
| QY | 301  | LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln   | 320  |
| Db | 3511 | CTCATCGACCAGTACCTGTATTACTTGAGCAGAGCAAAACACTCCAAGTGGAAACCACCACG | 3570 |
| QY | 321  | AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn   | 340  |
| Db | 3571 | CAGTCAAGGCTTCAGTTTCTCAGGCCGAGCGAGTGACATTCGGGACCAGTCTAGGAAC     | 3630 |
| QY | 341  | TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn   | 360  |
| Db | 3631 | TGGCTTCTTGACCCCTGTACCGCCAGCAGCGAGTATCAAAAGACATCTGCGGATAACAAC   | 3690 |
| QY | 361  | AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle   | 380  |
| Db | 3691 | AACAGTGAATATCTCGTGAAGTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTG  | 3750 |
| QY | 381  | IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet   | 400  |
| Db | 3751 | GTGAATCCGGGCCCCGCCATGGCAAGCCACAAGGACGATGAAGAAAAGTTTTTCTCTCAG   | 3810 |
| QY | 401  | SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn   | 420  |
| Db | 3811 | AGCGGGTTCATCTTTGGGAAGCAAGGCTCAGAGAAAACAAATGTGGACATTGAAAAG      | 3870 |
| QY | 421  | ValMetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe   | 440  |
| Db | 3871 | GTCTGATTACAGACGAGAGGAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTAT      | 3930 |
| QY | 441  | GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis   | 460  |
| Db | 3931 | GGTTCGTATCTACCAACCTCCAGAGAGGCAACAGACAGCAGCTACCGCAGATGTCAAC     | 3990 |
| QY | 461  | AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro   | 480  |
| Db | 3991 | ACACAAGGCGTTCTTCCAGGATGGTCTGGCAGGACAGAGATGTACTCTCAGGGGCC       | 4050 |
| QY | 481  | IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly   | 500  |
| Db | 4051 | ATCTGGGCAAGATTCCACACAGCGACGACATTTTTCACCCCTCTCCCTCATGGGTGGA     | 4110 |
| QY | 501  | PheGlyLeuLysAsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsn   | 520  |
| Db | 4111 | TTCCGACTTAAACACCCCTCCTCCACAGATTCTCATCAAGAACACCCCGGTACTGCGAAT   | 4170 |
| QY | 521  | ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly   | 540  |
| Db | 4171 | CCTTCGACCACCTTCAGTGGGCAAAAGTTTGTCTTCTCATCACAGTACTCCACGGGA      | 4230 |
| QY | 541  | GlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPro   | 560  |
| Db | 4231 | CAGGTCAGGTGGAGATCGAGTGGAGCTGCAGAGGAAACAGCAAAACGCTGGAATCCC      | 4290 |
| QY | 561  | GluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsn   | 580  |
| Db | 4291 | GAAATTCAGTACACTTCCAACTACAACAAGTCTGTTAATGTGGACTTTACTGTGGACACT   | 4350 |
| QY | 581  | AsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu      | 599  |
| Db | 4351 | AATGGCGTGTATTTCAGAGCCCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTG    | 4407 |

RESULT 5

US-08-254-358-1  
; Sequence 1, Application US/08254358  
; Patent No. 5658785  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Philip R.  
; TITLE OF INVENTION: Adeno-Associated Virus Materials and  
; TITLE OF INVENTION: Methods  
; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:  
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 S. Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/254,358  
; FILING DATE:  
; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5658785and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 31975  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4680 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-08-254-358-1

Alignment Scores:  
Pred. No.: 1.44e-253 Length: 4680  
Score: 2722.50 Matches: 492  
Percent Similarity: 89.00% Conservative: 42  
Best Local Similarity: 82.00% Mismatches: 64  
Query Match: 83.74% Indels: 2  
DB: 1 Gaps: 1

US-09-807-802A-15 (1-599) x US-08-254-358-1 (1-4680)

|    |      |  |      |
|----|------|--|------|
| Qy | 1    | ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer   | 20   |
| Db | 2614 | ACGGCTCCGGGAAAAAGAGCGGTAGAGCACTCTCTGTGGAGCCAGACTCTCTCTCG       | 2673 |
| Qy | 21   | GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly   | 40   |
| Db | 2674 | GGAACCGGAAGCGGGCCAGCGCTGCAAGAAAAGATTGAATTTGGTCAGACTGGA         | 2733 |
| Qy | 41   | AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla   | 60   |
| Db | 2734 | GACGCAGACTCAGTACCTTGACCCCGCCAGCTCTCGGACAGCCACCGAGCCCTCTGGT     | 2793 |
| Qy | 61   | ValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGly      | 80   |
| Db | 2794 | CTGGGAACCTAATACGATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGC   | 2853 |
| Qy | 81   | AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp   | 100  |
| Db | 2854 | GCCGACGGAGTGGGTAATTCCTCCGGAATTTGGCATTTGGATTCCACATGGATGGCGAC    | 2913 |
| Qy | 101  | ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr   | 120  |
| Db | 2914 | AGAGTCATACCACCAGCACCCGAACCTGGGCCCTGCCACCTTACAAACACCCCTCTAC     | 2973 |
| Qy | 121  | LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer   | 140  |
| Db | 2974 | AAACAAATTTCCAGCCCAATCA---GGAGCCTCGAACGACCAATCACTACTTTGGCTACAGC | 3030 |
| Qy | 141  | ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp   | 160  |

|    |      |   |      |
|----|------|---|------|
| Db | 3031 | ACCCCTTGGGGGTATTTTGACTTCAACAGATTCACCTGCCACTTTTCCACCACGTGACTGG | 3090 |
| Qy | 161  | GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe  | 180  |
| Db | 3091 | CAAAGACTCATCAACAACACTGGGATTCGACCCCAAGAGACTCAACTTCAAGCTCTTT    | 3150 |
| Qy | 181  | AsnIleGlnValLysGluValThrThrAsnAspGlyValThrIleAlaAsnAsnLeu     | 200  |
| Db | 3151 | AACATTCAAGTCAAAGAGGTCAACGAGAATGACGGTACGACGACGATTGCCAATAACCTT  | 3210 |
| Qy | 201  | ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer  | 220  |
| Db | 3211 | ACCAGCACGGTTCAAGTGTTTACTGACTCGAGTACCAGCTCCCGTACGTCTCTGGCTCG   | 3270 |
| Qy | 221  | AlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGly  | 240  |
| Db | 3271 | GCGCATCAAGGATGCCTCCCGCGTTTCCAGCAGACGCTTTCATGTTGTCACAGTATGGA   | 3330 |
| Qy | 241  | TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu  | 260  |
| Db | 3331 | TACCTCACCTGAACAACGCGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCCTGGAG   | 3390 |
| Qy | 261  | TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu  | 280  |
| Db | 3391 | TACTTTCTTCTCAGATGCTGCGTACCGGAACAACCTTTTACCTTACGTACACTTTTGAG   | 3450 |
| Qy | 281  | GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro  | 300  |
| Db | 3451 | GACGTTCTTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCT    | 3510 |
| Qy | 301  | LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln  | 320  |
| Db | 3511 | CTCATCGACCAGTACCTGTATTACTTGAGCAGAAACAACACTCCAAGTGGAAACCACACG  | 3570 |
| Qy | 321  | AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn  | 340  |
| Db | 3571 | CAGTCAAGGCTTCAGTTTCTCAGCGCGAGCGAGTGACATTCGGGACCGAGTCTAGGAAC   | 3630 |
| Qy | 341  | TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn  | 360  |
| Db | 3631 | TGGCTTCTCTGGACCCTGTTACCGCCAGCAGCGAGTATCAAGAGACATCTCGGATAACAAC | 3690 |
| Qy | 361  | AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle  | 380  |
| Db | 3691 | AACAGTGAATACTCGTGGACTGGAGTACCAAGTACCACCTCAATGGCAGAGACTCTCTG   | 3750 |
| Qy | 381  | IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet  | 400  |
| Db | 3751 | GTGAATCCGGGGCCGCCATGGCAAGCCCAAGGACGATGAAGAAAAGTTTTCCTCTCAG    | 3810 |
| Qy | 401  | SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn  | 420  |
| Db | 3811 | AGCGGGTTCTCATCTTTGGGAAGCAAGGCTCAGAGAAAAACAATGTGAACATTGAAAAG   | 3870 |
| Qy | 421  | ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe     | 440  |
| Db | 3871 | GTCAATGATTACAGACGAAGAGGAAATCGGAACAACCAATCCCGTGGCTACCGAGCAGTAT | 3930 |
| Qy | 441  | GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis  | 460  |
| Db | 3931 | GGTTCTGTATCTACCAACCTCCAGAGAGGCAACAGACAGCAGCTACCGCAGATGTCAAC   | 3990 |
| Qy | 461  | AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro  | 480  |
| Db | 3991 | ACACAAGGCGTTCTTCCAGGCATGGTCTGCGCAGACAGAGATGTGTACCTTCAGGGGCC   | 4050 |
| Qy | 481  | IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly  | 500  |
| Db | 4051 | ATCTGGCAAGATTCCACACCGGACGACATTTTCAACCTCTCCCTCTCATGGGTGGA      | 4110 |
| Qy | 501  | PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn     | 520  |
| Db | 4111 | TTCGGACTTAAACACCTCTCTCCACAGATTCTCATCAAGAAACACCCCGGTACCTGCGAAT | 4170 |



QY 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540  
|||  
Db 4171 CCTTCGACCACCTTCAGTGGGCAAGTTTGCTTCCTTCATCACACAGTACTCCACGGGA 4230  
QY 541 Gln-ValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPr 560  
|||  
Db 4231 CACGGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAAGGAAAACAGCAACCGCTGGAATCC 4290  
QY 560 oGluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAs 580  
|||  
Db 4291 CGAAATTCAGTACACTTCCAACTACAACAAGTCTGTTAATCGTGGACTTACCGTGGATAC 4350  
QY 580 nAsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599  
|||  
Db 4351 TAATGGCGTGATTTCAGAGCCCTCGCCCATTTGGCACCAATACCTGACTCGTAATCTG 4408

RESULT 6

US-08-475-391-1  
; Sequence 1, Application US/08475391  
; Patent No. 5786211  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Philip R.  
; TITLE OF INVENTION: Adeno-Associated Virus Materials and  
; TITLE OF INVENTION: Methods  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 S. Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/475,391  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/254,358  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 578621land, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 31975  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4680 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-475-391-1

Alignment Scores:  
Pred. No.: 1.44e-253 Length: 4680  
Score: 2722.50 Matches: 492  
Percent Similarity: 89.00% Conservative: 42  
Best Local Similarity: 82.00% Mismatches: 64  
Query Match: 83.74% Indels: 2  
DB: 1 Gaps: 1

US-09-807-802A-15 (1-599) x US-08-475-391-1 (1-4680)

QY 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSer 20

Db 2614 ACGGCTCCGGGAAAAAGAGCGGTAGAGCACTCTCCTGTGGAGCCAGACTCCTCTCG 2673  
QY 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40  
|||  
Db 2674 GGAACCGGAAAGCGGGCCAGCAGCCTGCAAGAAAAAGATTGAATTTGGTCAGACTGGA 2733  
QY 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60  
|||  
Db 2734 GACGCAGACTCAGTACCTGACCCCCAGCCTCTCGGACAGCCACCAGCAGCCCCCTCTG 2793  
QY 61 ValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnGluGly 80  
|||  
Db 2794 CTGGGAACATAATACGATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGC 2853  
QY 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100  
|||  
Db 2854 GCCGACGGAGTGGGTAATTCTCCGGAAATTGGCATTTCCACATGGATGGGCGAC 2913  
QY 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120  
|||  
Db 2914 AGAGTCATCACACCAGCACCCGAACTGGGCCCCCTGCCACCTACAAACACCTCTAC 2973  
QY 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140  
|||  
Db 2974 AAACAAATTTCCAGCCAATCA---GGAGCCTCGAACGACAATCACTACTTTGGCTACAGC 3030  
QY 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160  
|||  
Db 3031 ACCCTTGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCACCACGTGACTGG 3090  
QY 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180  
|||  
Db 3091 CAAAGACTCATCAACAACAACCTGGGGATTCGACCCCAAGAGACTCAACTTCAAGCTCTT 3150  
QY 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200  
|||  
Db 3151 AACATTCAAGTCAAAGAGGTACCGCAGAAATGACGGTACGACGAGATTGCCAATAACCTT 3210  
QY 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220  
|||  
Db 3211 ACCGACCGGTTCAGGTGTTTACTGACTCGGAGTACCAGTCCCCGTACGTCTCGGCTCG 3270  
QY 221 AlaHisGlnGlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGly 240  
|||  
Db 3271 GCGCATCAAGGATGCTCCCGCGTTCGCCAGCAGACGTCTTTCATGGTGCCACAGTATGGA 3330  
QY 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260  
|||  
Db 3331 TACCTCACCCCTGAACAACGGGAGTCAGGCAGTAGGACGCTCTTCAATTTTACTGCTTGAG 3390  
QY 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280  
|||  
Db 3391 TACTTTCCCTTCTCAGATGCTGCGTACCGGAACAACACTTACCTTCAGCTACACTTTTGAG 3450  
QY 281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300  
|||  
Db 3451 GACGTTCCCTTTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGCTCATGAATCCT 3510  
QY 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320  
|||  
Db 3511 CTCATCGACCAGTACCTGTATTACTTGTAGCAGAGAAACAAACACTCCAAAGTGAACACCACG 3570  
QY 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340  
|||  
Db 3571 CAGTCAAGGCTTCAGTTTCTCAGGCCGAGCGAGTGACATTCGGGACCAGTCTAGGAAC 3630  
QY 341 TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360  
|||  
Db 3631 TGGCTTCTCTGGACCCCTGTTACCGCCAGCAGGAGTATCAAAGACATCTGCGGATAACAC 3690  
QY 361 AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380  
|||

Db 3691 AACAGTGAATACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTG 3750  
Qy 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400  
Db 3751 GTGAATCCGGGGCCCGCCATGGCAAGCCACAAAGGACGATGAAGAAAGTTTTTTCCTCAG 3810  
Qy 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420  
Db 3811 AGCGGGTTCATCTTTGGGAAGCAAGGCTCAGAGAAACAAATGTGAACATTGAAAG 3870  
Qy 421 ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440  
Db 3871 GTCATGATTACAGACGAAGAGGAATCGGAACCAACCAATCCCGTGGCTACGGAGCAGTAT 3930  
Qy 441 GlyThrValAlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHis 460  
Db 3931 GGTTCGTATCTACCAACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAAC 3990  
Qy 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTrpLeuGlnGlyPro 480  
Db 3991 ACACAAGGCTTCTCCAGGCATGTTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCC 4050  
Qy 481 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500  
Db 4051 ATCTGGGCAAGATTCCACACACGGACGGACATTTTCACCCCTCTCCCTCATGGGTGA 4110  
Qy 501 PheGlyLeuLysAsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520  
Db 4111 TTCGGACTTAAACACCTCTCCACAGATCTCATCAAGAACACCCCGTACCTGCGAAT 4170  
Qy 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540  
Db 4171 CCTTCGACCCTTCAGTGGCGCAAGTTTGCTTCTTCATCACAGTACTCCACGGGA 4230  
Qy 541 Gln-ValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPr 560  
Db 4231 CACGGTCAGCGTGAGATCGAGTGGAGCTGCAGAAAGGAAACAGCAACGCTGGAATCC 4290  
Qy 560 oGluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAs 580  
Db 4291 CGAAATTCAGTACACTTCCAACTACAACAAGTCTGTTAATCGTGGACTTACCGTGATAC 4350  
Qy 580 nAsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599  
Db 4351 TAATGGCGTGATTACAGAGCCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTG 4408

RESULT 7  
US-08-709-609-1

; Sequence 1, Application US/08709609  
; Patent No. 5858775  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Philip R.  
; TITLE OF INVENTION: Adeno-Associated Virus Materials and  
; TITLE OF INVENTION: Methods  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 S. Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/709,609  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5858775and, Greta E.

; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 31975  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4680 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-08-709-609-1  
  
Alignment Scores:  
Pred. No.: 1.44e-253 Length: 4680  
Score: 2722.50 Matches: 492  
Percent Similarity: 89.00% Conservative: 42  
Best Local Similarity: 82.00% Mismatches: 64  
Query Match: 83.74% Indels: 2  
DB: 1  
  
US-09-807-802A-15 (1-599) x US-08-709-609-1 (1-4680).  
  
Qy 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20  
Db 2614 ACGGCTCCGGGAAAAAGAGGCCGGTAGAGCACTCTCCTGTGGAGCCAGACTCCTCTCG 2673  
Qy 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40  
Db 2674 GGAACCGGAAAGCGCGGCAGCAGCCTGCAAGAAAAAGATTGAATTTTGGTCAGACTGGA 2733  
Qy 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60  
Db 2734 GACGCACTCAGTACTGACCCCGCAGCCTCTCGGACAGCCACGAGCCCTCTGCT 2793  
Qy 61 ValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGly 80  
Db 2794 CTGGAACTAATACGATGGTACAGGCAGTGCGGCACCAATGGCAGACAATAACGAGGCG 2853  
Qy 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100  
Db 2854 GCCGCGGAGTGGTAAATCTCCGGAATTTGGCATTTCCACATGGATGGGCGGAC 2913  
Qy 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120  
Db 2914 AGAGTCATCACCACCAGCACCGCAACCTGGGCCCTGCCACCTACAACACCCTCTAC 2973  
Qy 121 LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140  
Db 2974 AAACAAATTTCCAGCCCAATCA--GGAGCCTCGAACGACAAATCACTACTTTGGCTACAGC 3030  
Qy 141 ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160  
Db 3031 ACCCTTGGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCCACACGTGACTGG 3090  
Qy 161 GlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180  
Db 3091 CAAGACTCATCAACAACAACACTGGGATTCGACCCCAAGAGACTCAACTTCAAGCTCTTT 3150  
Qy 181 AsnIleGlnValLysGluValThrThrAsnAspGlyValThrIleAlaAsnAsnLeu 200  
Db 3151 AACATTCAAGTCAAAGAGGTACGCGAATGACGGTACGACGAGATTGCCAATAACCTT 3210  
Qy 201 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220  
Db 3211 ACCAGCAGGTTTCAAGTGTGTTACTGACTCGGAGTACAGCTCCCGTACGTCTCGGCTCG 3270  
Qy 221 AlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyrGly 240  
Db 3271 GCGCATCAAGGATGCTCTCCCGCGTTCCTCCAGCAGACGTCCTTCTCATGGTGCCACAGTATGGA 3330

Qy 241 TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260  
Db 3331 TACCTCACCTGAACAACGGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCTCGGAG 3390  
Qy 261 TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280  
Db 3391 TACTTCTCTCAGATGCTGCTACCGGAACAACATTTCCTTCAGCTACACTTTTGAG 3450  
Qy 281 GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300  
Db 3451 GACGTCTCTTCCACAGCAGCTACGCTCACAGCCAGAGCTCGGACCGTCTCATGAATCCT 3510  
Qy 301 LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320  
Db 3511 CTCATCAGCAGTACCTGTATTACTTGAGCAGACAACAACACTCCAAGTGAACCCACCG 3570  
Qy 321 AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340  
Db 3571 CAGTCAAGGCTTCAGTTTCTCAGGCGGAGCGAGTGACATTCGGGACCGTCTAGGAAC 3630  
Qy 341 TrpLeuProGlyProCysTyrArgGlnArgValSerLysThrLysThrAspAsnAsn 360  
Db 3631 TGGCTTCCTGGACCCCTGTTACCGCCAGCAGCGAGTATCAAGACATCTCGGATAACAAC 3690  
Qy 361 AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380  
Db 3691 AACAGTGAATACTCGTGGACTGGAGTACCAAGTACCACCTCAATGGCAGAGACTCTCTG 3750  
Qy 381 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400  
Db 3751 GTGAATCCGGGCGCCGCCATGGCAAGCCACAAGGACGATGAAGAAAAGTTTTTCTCTCAG 3810  
Qy 401 SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420  
Db 3811 AGCGGGGTCTCTCATCTTTGGGAAGCAAGGCTCAGAGAAAACAATGTGAACATTTGAAAG 3870  
Qy 421 ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440  
Db 3871 GTCATGATTACAGACGAGAGGAATCGGAACAACCAATCCCGTGGCTACGGAGCAGTAT 3930  
Qy 441 GlyThrValAlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHis 460  
Db 3931 GGTTCTGTATCTACCAACCTCCAGAGAGGCAACAGACAGCTACCGCAGATGTCAAC 3990  
Qy 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 480  
Db 3991 ACACAAGGCGTTCTTCCAGGCATGGTCTGGCAGGACAGAGATGTACCTTCAGGGGCC 4050  
Qy 481 IleTrpAlaIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500  
Db 4051 ATCTGGGCAAGATTCCACACACGACGAGCAGATTTTCACCCCTCTCCCTCATGGGTGA 4110  
Qy 501 PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520  
Db 4111 TTCGGACTTAAACACCCCTCTCCACAGATTCTCATCAAGAACAACCCCGGTACCTGCGAAT 4170  
Qy 521 ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540  
Db 4171 CCTTCGACCCACTTCAGTGGCGCAAGTTTGTCTTCTCATCACACAGTACTCCACGGGA 4230  
Qy 541 Gln-ValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPr 560  
Db 4231 CACGGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAAAACAGCAACCGTGGAAATCC 4290  
Qy 560 oGluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAs 580  
Db 4291 CGAAATTCAGTACACTTCCAACACTACAACAGTCTGTAAATCGTGGACTTACCGTGATAC 4350  
Qy 580 nAsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599  
Db 4351 TAATGGCGGTGATTTCAGAGCCTCGCCCAATTGGCACCAGATACCTGACTCGTAATCTG 4408  
RESULT 8

PCT-US95-07178-1  
; Sequence 1, Application PC/TUS9507178  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Philip R.  
; TITLE OF INVENTION: Adeno-Associated Virus Materials and  
; TITLE OF INVENTION: Methods  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 S. Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/07178  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Noland, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 31975  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4680 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
PCT-US95-07178-1

Alignment Scores:  
Pred. No.: 1.44e-253 Length: 4680  
Score: 2722.50 Matches: 492  
Percent Similarity: 89.00% Conservative: 42  
Best Local Similarity: 82.00% Mismatches: 64  
Query Match: 83.74% Indels: 2  
DB: 5 Gaps: 1

US-09-807-802A-15 (1-599) x PCT-US95-07178-1 (1-4680)

Qy 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20  
Db 2614 ACGGCTCCGGGAAAAAGAGCGCGGTAGAGCACTCTCTGTGGAGCCAGACTCTCTCTCG 2673  
Qy 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPheGlyGlnThrGly 40  
Db 2674 GGAACCGGAAGCGCGGCCAGCAGCTGCAAGAAAAAGATTGAATTTTGGTCAGACTGGA 2733  
Qy 41 AspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAla 60  
Db 2734 GACGCACTCAGTACCTGTACCTCCCGCCAGCCTCTCGGACAGCCAGCCCTCTCTGT 2793  
Qy 61 ValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGly 80  
Db 2794 CTGGAACTAATACGATGGGTGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGG 2853  
Qy 81 AlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAsp 100  
Db 2854 GCCGACGGAGTGGGTAAATCTCTCCGGAATTTGGCATTTCCATGGATGGGCGAC 2913  
Qy 101 ArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyr 120  
Db 2914 AGAGTCATCACCACCGACCGCAACCTGGGCCCTGGCCCACTACCAACCACTCTCTAC 2973



```

QY      121  LysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSer 140
Db      2974 AAACAAATTTCAGCCCAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGC 3030

QY      141  ThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrp 160
Db      3031 ACCCTTGGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCACCCACGTGACTGG 3090

QY      161  GlnArgIleuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPhe 180
Db      3091 CAAAGACTCATCAACAACAACATGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTT 3150

QY      181  AsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 200
Db      3151 AACATTCAAGTCAAAAGAGGTACCGAGAATGACGGTACGACGACGATTGCCAATAACCTT 3210

QY      201  ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 220
Db      3211 ACCAGCAGGTTTCAGGTGTTTACTGACTCGGAGTACCAGTCCCGTACGTCTCTCGGCTCG 3270

QY      221  AlaHisGlnGlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGly 240
Db      3271 GCGCATCAAGGATGCTCCCGCGTTCCAGAGCTCCAGAGACGTCTTCATGGTGCACAGTATGA 3330

QY      241  TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu 260
Db      3331 TACCTCACCTGAACAACGGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCTGGAG 3390

QY      261  TyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu 280
Db      3391 TACTTTCTTCTCAGATGCTGCGTACCGGAACAACATTTACCTTCAGCTACACTTTTGAG 3450

QY      281  GluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnPro 300
Db      3451 GACGTTCTTTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCT 3510

QY      301  LeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGln 320
Db      3511 CTCATCGACCAGTACCTGTATTACTTGAGCAGAAACAACACTCCAAGTGAACCCACCG 3570

QY      321  AsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsn 340
Db      3571 CAGTCAAGGCTTCAGTTTCTCAGGCCGAGCGAGTGACATTCGGGACCAGTCTAGGAAC 3630

QY      341  TrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 360
Db      3631 TGGCTTCCTGGACCCCTGTTACCGCCAGCAGCGAGTATCAAAGACATCTGCGGATAACAAC 3690

QY      361  AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 380
Db      3691 AACAGTGAATACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGCGAGAGACTCTCTG 3750

QY      381  IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 400
Db      3751 GTGAATCCGGGGCCCGCCATGGCAAGCCACAAGGACGATGAAGAAAAGTTTTTTCCTCAG 3810

QY      401  SerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsn 420
Db      3811 AGCGGGGTTCTCATCTTTGGGAAGCAAGGCTCAGAGAAAACAATGTGAACATTGAAAAG 3870

QY      421  ValMetIleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPhe 440
Db      3871 GTCATGATTACAGACGAAGAGGAAATCGGAACAACCAATCCCGTGGCTACGGAGCAGTAT 3930

QY      441  GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis 460
Db      3931 GGTTCGTATCTACCAACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAAC 3990

QY      461  AlaMetGlyAlaLeuProGlyMetValTyrGlnAspArgAspValTyrLeuGlnGlyPro 480
Db      3991 ACACAAGGCGTTCTTCCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCC 4050

```

```

QY      481  IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 500
Db      4051 ATCTGGGCAAGATTCCACACACGGACGGACATTTTCACCCCTCTCCCCTCATGGGTGGA 4110

QY      501  PheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsn 520
Db      4111 TTCGGACTTAAACACCCCTCTCCACAGATTCTCATCAAGAACACCCCGGTACTGCGAAT 4170

QY      521  ProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGly 540
Db      4171 CCTTCGACCACCTTCAGTGGCGCAAAGTTTGTTCTTCATCACACAGTACTCCACGGGA 4230

QY      541  Gln-ValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnPr 560
Db      4231 CACGGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAAGGAACAGCAAACGCTGGAATCC 4290

QY      560  oGluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAs 580
Db      4291 CGAAATTTCAGTACACTTCCAACACTACAACAAGTCTGTTAATCGTGGACTTACCGTGGATAC 4350

QY      580  nAsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599
Db      4351 TAATGGCGTGATTTCAGAGCCTCGCCCCATTGGCACCAGATACCTGACTCGTAATCTG 4408

RESULT 9
US-09-438-268-2
; Sequence 2, Application US/09438268
; Patent No. 6491907
; GENERAL INFORMATION:
; APPLICANT: Rabinowitz, Joseph E.
; APPLICANT: Samulski, Richard J
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 5470-186
; CURRENT APPLICATION NUMBER: US/09/438,268
; CURRENT FILING DATE: 1999-11-10
; EARLIER APPLICATION NUMBER: 60/107,840
; EARLIER FILING DATE: 1998-11-10
; EARLIER APPLICATION NUMBER: 60/123,651
; EARLIER FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 8151
; TYPE: DNA
; ORGANISM: Virus
US-09-438-268-2

```

```

Alignment Scores:
Pred. No.:      9.23e-222      Length:      8151
Score:          2398.50      Matches:      435
Percent Similarity: 82.17%      Conservative: 58
Best Local Similarity: 72.50%      Mismatches: 96
Query Match:      73.78%      Indels:      11
DB:              4              Gaps:      3

```

US-09-807-802A-15 (1-599) x US-09-438-268-2 (1-8151)

```

QY      1  ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSer 20
Db      2544 ACGGCTCCTGGAAAGAAGAGAGACCGTTGATTGAATCCCCCCCAGCAGCCGACTCCTCCACG 2603

QY      21  GlyTleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPhe---GlyGlnThr 39
Db      2604 GGTATCGGCAAAAAAGGCAAGCAGCCGGCTAAAAAAGAGCTCGTTTTCGAAGACGAAACT 2663

QY      40  GlyAspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAla 59
Db      2664 GGAGCAGCGCAGCGACCCCTGAGGATCAACTTCCGGA-----GCC 2705

QY      60  AlaValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGlu 79

```

Db 2706 ATGTCTGATGACAGTGAGATGCGTGCAGCAGCTGGCGGAGCTGCAGTCGAGGGCGGACAA 2765

Qy 80 GlyAlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGly 99

Db 2766 GGTGCCGATGGAGTGGGTAATGCTCGGTTGATTGGCATTTGGCATTCACCTCGTCTGAG 2825

Qy 100 AspArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeu 119

Db 2826 GGCCACGTACGACCACCCAGCAGCAGAACCTGGGTCTTGCCACCTACAAACACCACTC 2885

Qy 120 TyrLysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyr 139

Db 2886 TACAAGCGACTCGGAGAG-----AGCCTGCAGTCCAACACCTACAAACGGATTTC 2933

Qy 140 SerThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAsp 159

Db 2934 TCCACCCCTGGGGATACTTTGACTTCAACCGCTTCCACTGCCACTTCTCACCAACGTGAC 2993

Qy 160 TrpGlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeu 179

Db 2994 TGGCAGCGACTCATCAACAACAACTGGGGCATGCGACCCCAAAGCCATGCGGGTCAAAATC 3053

Qy 180 PheAsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsn 199

Db 3054 TTCAACATCCAGGTCAAGGAGGTACGACGTCCGAACGGCGAGACAACGGTGGCTAATAAC 3113

Qy 200 LeuThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGly 219

Db 3114 CTTACCAGCACGGTTTCAGATCTTTGGGACTCGTCGTACGAACGTCCCGTACGTCTCTCGGC 3173

Qy 220 SerAlaHisGlnGlyCysLeuProPropheProAlaAspValPheMetIleProGlnTyr 239

Db 3174 TCGGCGCATCAAGGATGCCTCCCGCGTTCCCGAGCAGCGTCTTCATGGTCCACAGTAT 3233

Qy 240 GlyTyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeu 259

Db 3234 GGATACCTCACCCCTGAACAACGGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCGTG 3293

Qy 260 GluTyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPhe 279

Db 3294 GAGTACTTTCCTTCTCAGATGCTGCTACCGGAAACAACCTTTACCTTCAGCTACACTTT 3353

Qy 280 GluGluValPropheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsn 299

Db 3354 GAGGACGTTCTCTTCCACAGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAAT 3413

Qy 300 ProLeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAla 319

Db 3414 CCTCTCATCGACCAGTACCTGTATTACTTGAGCAGACAACAACACTCCAAGTGGAACCA 3473

Qy 320 GlnAsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLys 339

Db 3474 ACGCAGTCAAGGCTTCAGTTTCTCAGCGCGGAGCGAGTGACATTGGGACCAGTCTAGG 3533

Qy 340 AsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsn 359

Db 3534 AACTGGCTTCTGGACCCCTGTTACCGCCAGCAGCGAGTATCAAAGACATCTGCGGATAAC 3593

Qy 360 AsnAsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSer 379

Db 3594 AACAACTGAATACTCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCT 3653

Qy 380 IleIleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePro 399

Db 3654 CTGGTGAATCCGGGCCCGCCCATGGCAAGCCACAAGGACGATGAAGAAAAGTTTTTCT 3713

Qy 400 MetSerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAsp 419

Db 3714 CAGAGCGGGTCTCATCTTTGGGAACCAAGGCTCAGAGAAAACAATGTGAACATTGAA 3773

Qy 420 AsnValMetIleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArg 439

Db 3774 AAGGTCATGATTACAGACGAAGAGGAATCGGAACAACCAATCCCGTGGCTACGGAGCAG 3833

Qy 440 PheGlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspVal 459

Db 3834 TATGGTTCTGTATCTACCAACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTC 3893

Qy 460 HisAlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGly 479

Db 3894 AACACACAAGCGTTCTTCCAGGCATGGTCTGGCAGGACAGAGATGTGTACTTCAGGGG 3953

Qy 480 ProIleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGly 499

Db 3954 CCCATCTGGGCAAGATTCCACACACGCGGACATTTTCACCCCTCTCCCTCATGGGT 4013

Qy 500 GlyPheGlyLeuLysAsnProProProGlnIleLeuIleLysAsnThrProValProAla 519

Db 4014 GGATTTCGGACTTAAACACCCCTCCTCCACAGATTCTCATCAAGAACACCCCGGTACTCGG 4073

Qy 520 AsnProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThr 539

Db 4074 AATCCTTCGACCACCTTCAGTGCAGCAAGTTTGTCTTCTTCATCACACAGTACTCCACG 4133

Qy 540 GlyGlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsn 559

Db 4134 GGACAGGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAAGGAAACAGCAAACGCTGGAAT 4193

Qy 560 ProGluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAsp 579

Db 4194 CCCGAAATTCAGTACACTTCCAACACTACAACAAGTCTGTTAATCGTGGACTTACCGTGGAT 4253

Qy 580 AsnAsnGlyLeuTyrThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599

Db 4254 ACTAATGGCGTGTATTTCAGAGCCTCGCCCCATTGGCACCAAGATACCTGACTCGTAATCTG 4313

RESULT 10

US-09-532-594B-17

; Sequence 17, Application US/09532594B

; Patent No. 6468524

; GENERAL INFORMATION:

; APPLICANT: Chorini, John A.

; APPLICANT: Kotin, Robert M.

; APPLICANT: Safer, Brian

; APPLICANT: Davidson, Beverly

; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF

; FILE REFERENCE: 14014.0252U2

; CURRENT APPLICATION NUMBER: US/09/532,594B

; CURRENT FILING DATE: 2000-03-22

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 17

; LENGTH: 1800

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =

; OTHER INFORMATION: synthetic construct

; NAME/KEY: misc\_feature

; LOCATION: 342

; OTHER INFORMATION: n = g, a, c or t(u)

; NAME/KEY: misc\_feature

; OTHER INFORMATION: AAV4 capsid protein VP2 gene

US-09-532-594B-17

Alignment Scores:

|                        |         |               |      |
|------------------------|---------|---------------|------|
| Pred. No.:             | 1e-167  | Length:       | 1800 |
| Score:                 | 1830.50 | Matches:      | 355  |
| Percent Similarity:    | 69.87%  | Conservative: | 74   |
| Best Local Similarity: | 57.82%  | Mismatches:   | 154  |
| Query Match:           | 56.31%  | Indels:       | 31   |
| DB:                    | 4       | Gaps:         | 10   |

US-09-807-802A-15 (1-599) x US-09-532-594B-17 (1-1800)

Qy 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20

|    |     |  |     |
|----|-----|--|-----|
| Db | 1   | ACGGCTCCTGGAAAGAAGAGACCGTTGATTGAATCCCCCCAGCAGCCGAGCTCCTCCACG   | 60  |
| Qy | 21  | GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPhe--GlyGlnThr    | 39  |
| Db | 61  | GGTATCGGCATAAAGGCAAGCAGCCGGCTAAAGAGAGCTCGTGTTCGAAGACGAAACT     | 120 |
| Qy | 40  | GlyAspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAla   | 59  |
| Db | 121 | GGAGCAGCGGACCGGACCCCTGAGGGATCAACTTCGGA-----GCC                 | 162 |
| Qy | 60  | AlaValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGlu      | 79  |
| Db | 163 | ATGCTGATGACAGTGAGATGCGTGACGAGCTGGCGGAGTGCAGTCGAGGSGGACAA       | 222 |
| Qy | 80  | GlyAlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGly   | 99  |
| Db | 223 | GGTCCGATGGAGTGGTAAATGCCTCGGTGATTGGCATTTGGATTCCACCTGGTCTGAG     | 282 |
| Qy | 100 | AspArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeu   | 119 |
| Db | 283 | GGCCACGTCACGACCACGACGACCAACCTGGGTCTTGCCCACTCAACAACCCACTN       | 342 |
| Qy | 120 | TyrLysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyr   | 139 |
| Db | 343 | TACAAGCGACTCGGAGAG-----AGCCTGCAGTCCAAACACCTACAACCGGATTC        | 390 |
| Qy | 140 | SerThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAsp   | 159 |
| Db | 391 | TCCACCCCTGGGATACTTTGACTTCAACCGCTTCCACTGCCACTTCTCACCACGTTGAC    | 450 |
| Qy | 160 | TrpGlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeu   | 179 |
| Db | 451 | TGGCAGCGACTCATCAACAACAACACTGGGCGCATGCGACCCAAAGCCATCGGGTCAAAATC | 510 |
| Qy | 180 | PheAsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsn   | 199 |
| Db | 511 | TTCAACATCCAGGTCAAGGAGGTACGACGTCGAACCGCGAGACACGGTGGCTAATAAC     | 570 |
| Qy | 200 | LeuThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGly   | 219 |
| Db | 571 | CTTACCAGCAGGTTTCAGATCTTTGCGGACTCGTCGTACGAACTGCCGTACGTGATGGAT   | 630 |
| Qy | 220 | SerAlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyr   | 239 |
| Db | 631 | GCGGGTCAAGAGGGCGAGCCTGCCTCTTTCCCAACGACGCTCTTATGTGTCCTCCAGTAC   | 690 |
| Qy | 240 | GlyTyr-----LeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPhe       | 256 |
| Db | 691 | GGCTACTGTGAGACTGGTGACCGGCAACACTTCGCAGCAACAGACTGACAGAAATGCCTTC  | 750 |
| Qy | 257 | TyrCysLeuGluTyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSer   | 276 |
| Db | 751 | TACTGCCTGGAGTACTTTCTTCGCAGATGCTGCGGACTGGCAACAACATTGAAATACG     | 810 |
| Qy | 277 | TyrThrPheGluGluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArg   | 296 |
| Db | 811 | TACAGTTTGAGAAGGTGCTTTTCCACTCGATGTCGCGCACAGCCAGAGCTGGACCGG      | 870 |
| Qy | 297 | LeuMetAsnProLeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsn-----    | 314 |
| Db | 871 | CTGATGAACCTCTCATCGACCAGTACTGTGGGACTGCAATCGACCCACCGGAACC        | 930 |
| Qy | 315 | -----GlnSerGlySerAlaGlnAsnLysAspLeuLeuPheSerArgGlySerProAla    | 332 |
| Db | 931 | ACCTGAATGCGGGGACTGCCACCACCAAC-----TTTACCAGCTCGCGCCTACC         | 981 |
| Qy | 333 | GlyMetSerValGlnProLysAsnTrpLeuProGlyProCysTyrArgGlnGlnArgVal   | 352 |
| Db | 982 | AACCTTTTCCAACCTTAAAGAAACCTGGCTGCCCGGGCCTTCAATCAAGCAGCAGGGCTTC  | 104 |
| Qy | 353 | SerLysThrLysThrAspAsnAsnAsnSerAsnPheThrTrpThrGlyAlaSer-----    | 370 |

```

Db      1042 TCAAGACTGCC-----AATCAAAACTACAAGATCCCTGTCCACCGGGTCAGACAGTCTC 1099
QY      371  ---LysTyr-----AsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 385
Db      1096 ATCAAAATACGACGACGACAGCACTCTGGACGGAAGATGGAGTGCCTCGACCCCGGACCT 1155
QY      386 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 405
Db      1156 CCAATGGCCACGGCTGGACCTGCGGACAGCAAG--TTCAGCAACAGCCAGCTCATCTTT 1212
QY      406 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 425
Db      1213 GCGGGGCCCTAAACAGAACGGCAACACGGCCACCGTACCCGGGACTCTGTATCTTCACCTCT 1272
QY      426 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 445
Db      1273 GAGGAGGAGTGGCAGCCACCAACGCCACCGATACGGACATGTGGGGCAACCTACCTGGC 1332
QY      446 AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 465
Db      1333 GGTGACCAGAGCAACAGCAACCTGCCGACCGTGGACAGACTGACAGCCTTGGAGCCCGTG 1392
QY      466 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 485
Db      1393 CCTGGAATGCTGGCAAAACAGAGACATTACTACAGGGTCCCATTGTGGGCAAGATT 1452
QY      486 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 505
Db      1453 CCTCATACCGATGGACACTTTCACCCCTCACCGCTGATGGTGGGTTGGGCTGAAACAC 1512
QY      506 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 525
Db      1513 CCGCCTCTCTCAATTTTATCAAGAACACCCCGGTACCTGCGAATCCTGCAACGACCTTC 1572
QY      526 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 545
Db      1573 AGCTCTACTCCGGTAAACTCCTTCAATTAAGAACACCCCGGTACCTGCGCAGGTGTCCGTGCAG 1632
QY      546 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 565
Db      1633 ATTGACTGGGAGATCCAGAAAGGAGCGGTCCAAACGCTGGAAACCCCGAGGTCCAGTTTACC 1692
QY      566 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 585
Db      1693 TCCAACCTACGGACAGCAAAACTCTCTGTGTGGGCTCCCGATGCGGCTGGGAAATACACT 1752
QY      586 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599
Db      1753 GAGCCTAGGGCTATCGGTACCCGCTACCTCACCACCACCTG 1794

RESULT 11
US-09-532-594B-5
; Sequence 5, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2208
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: synthetic construct
; NAME/KEY: misc feature

```

setl y s i n t l y a i n t a b r a s i n a m b i s e t a v e n e m i t t i r n i n g c y a t a t  
:::



```
; LOCATION: 750
; OTHER INFORMATION: n = g, a, c or t(u)
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 caspid protein VP1 gene
US-09-532-594B-5

Alignment Scores:
Pred. No.:      1.39e-167      Length:      2208
Score:          1830.50        Matches:      355
Percent Similarity: 69.87%      Conservative:  74
Best Local Similarity: 57.82%    Mismatches:   154
Query Match:     56.31%        Indels:       31
DB:              4            Gaps:         10

US-09-807-802A-15 (1-599) x US-09-532-594B-5 (1-2208)

QY      1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20
Db      409 ACGGCTCCTGGAAGAGAGACCGTTGATTGAATCCCCCCAGCAGCCCGACTCCTCCACG 468

QY      21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPhe---GlyGlnThr 39
Db      469 GGTATCGGCAAAAAGGCAAGCAGCCGGCTAAAAAGAAAGCTCGTTTTCGAAGACGAAACT 528

QY      40 GlyAspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAla 59
Db      529 GGAGCAGCGCAGCGACCCCTGAGGGATCAACTTCCGGA-----GCC 570

QY      60 AlaValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGlu 79
Db      571 ATGTCGTGATGACAGTGAGATGCGTCGACGAGCTGGCGAGCTGCAGTCGAGGSGGACAA 630

QY      80 GlyAlaAspGlyValGlyAsnAlaSerGlyAsnTyrHisCysAspSerThrTyrLeuGly 99
Db      631 GGTGCGGATGAGTGGTAATGCTCGGTGATTGGCATTCGATTCCACCTGGTCTGAG 690

QY      100 AspArgValIleThrThrSerThrArgThrTyrAlaLeuProThrTyrAsnAsnHisLeu 119
Db      691 GGCCACGTCACGACCACCAAGCAGCAACCTGGGTCTTGCCACCTACAACAACCACTN 750

QY      120 TyrLysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyr 139
Db      751 TACAAGCGACTCGGAGAG-----AGCCTGCAGTCCAACACCTACAACCGGATTC 798

QY      140 SerThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAsp 159
Db      799 TCCACCCCTGGGATACTTTGACTTCAACCGCTTCCACTGCCACTTCTCAACCACGTCAC 858

QY      160 TrpGlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeu 179
Db      859 TGGCAGCGACTCATCAACAACAACCTGGGGCATGCGACCCCAAGCCATGCGGGTCAAAATC 918

QY      180 PheAsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsn 199
Db      919 TTCAACATCCAGGTCAAGGAGGTTCAGACGTGCGAACCGCGAGACAACGGTGGCTAATAAC 978

QY      200 LeuThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGly 219
Db      979 CTTACCAGCAGCGTTTCAGATCTTTGGGACTCGTCGTACGAAGTCCCGTACGTGATGGAT 1038

QY      220 SerAlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyr 239
Db      1039 GCGGGTCAAGAGGGCAGCCTGCCTCCTTTTCCCAACGACGCTCTTATGGTCCCCAGTAC 1098

QY      240 GlyTyr-----LeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPhe 256
Db      1099 GGCTACTGTGAGTGGTGACCGGCAACACTTCGAGCNAACAGACTGACAGAAATGCCTTC 1158

QY      257 TyrCysLeuGluTyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSer 276
Db      1159 TACTGCTGGAGTACTTTCCTTCGCAGATGCTGCGGACTGGCAACAACACTTTGAAATTACG 1218

QY      277 TyrThrPheGluGluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArg 296
```

```
Db      1219 TACAGTTTGAAGAGTGCCTTTTCCACTCGATGTACGGCACAGCCAGAGCCTGGACCGG 1278

QY      297 LeuMetAsnProLeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsn----- 314
Db      1279 CTGATGAACCCCTCTCATCGACCAGTACCTGTGGGACTGCAATCGACCACCCGGAACC 1338

QY      315 -----GlnSerGlySerAlaGlnAsnLysAspLeuPheSerArgGlySerProAla 332
Db      1339 ACCCTGAATGCCGGGACTGCCACCACCAAC-----TTTACCAAGCTGCGGCCTACC 1389

QY      333 GlyMetSerValGlnProLysAsnTrpLeuProGlyProCysTyrArgGlnArgVal 352
Db      1390 AACTTTTCCAACCTTAAAGAACTGGCTGCCCGGGCCTTCAATCAAGCAGCAGGGCTTC 1449

QY      353 SerLysThrLysThrAspAsnAsnAsnSerAsnPheThrTrpThrGlyAlaSer----- 370
Db      1450 TCAAAGACTGCC-----AATCAAAACTACAAGATCCCTGCCACCGGTCCAGACAGTCTC 1503

QY      371 ---LysTyr-----AsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 385
Db      1504 ATCAAAATACGAGACGCACAGCACTCTGGACGGAAGATGGAGTGCCTGACCCCGGACCT 1563

QY      386 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 405
Db      1564 CCAATGGCCACGCTGGACCTGCGGACAGCAAG--TTCAGCAACAGCCAGCTCATCTTT 1620

QY      406 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 425
Db      1621 GCGGGGCTAAACAGAAACGGCAACACCGGCCACCGTACCCGGGACTCTGTATCTTCACCTCT 1680

QY      426 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 445
Db      1681 GAGGAGGAGCTGGCAGCCACCAACGCCACCGATACGGACATGTGGGGCAACCTACCTGGC 1740

QY      446 AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 465
Db      1741 GGTGACCAGAGCAACAGCAACCTGCCGACCGTGGACAGACTGACAGCCTTGGGAGCCGTG 1800

QY      466 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 485
Db      1801 CCTGGAATGGTCTGCAAAAACAGAGACATTTACTACAGGGTCCCATTTGGGCCAAGATT 1860

QY      486 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 505
Db      1861 CCTCATACCGATGGACACTTTTCAACCCCTCACCGCTGATTGGTGGTTTGGGCTGAAACAC 1920

QY      506 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 525
Db      1921 CCGCCTCCTCAAATTTTATCAAGAAACACCCCGGTACCTGCGAATCCTGCAACGACCTTC 1980

QY      526 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 545
Db      1981 AGCTCTACTCCGGTAAACTCCTTTCATTACTCAGTACAGCACTGGCCAGGTGTCGGTGCAG 2040

QY      546 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 565
Db      2041 ATTGACTGGGAGATCCAGAAGGAGCGGTCCAACCGCTGGAACCCCGAGGTCCAGTTACC 2100

QY      566 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 585
Db      2101 TCCAACACTACGACAGCAAAAACCTCTCTGTGTGGGCTCCCGATGCGGTGGGAATAACT 2160

QY      586 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599
Db      2161 GAGCCTAGGGCTATCGGTACCCGCTACCTCACCACCACCTG 2202
```

RESULT 12  
US-09-532-594B-1  
; Sequence 1, Application US/09532594B  
; Patent No. 6468524  
; GENERAL INFORMATION:  
; APPLICANT: Chorini, John A.

```
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4767
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; NAME/KEY: misc_feature
; LOCATION: 3009
; OTHER INFORMATION: n = g, a, c or t (u)
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 genome
; US-09-532-594B-1

Alignment Scores:
Pred. No.: 4,75e-167 Length: 4767
Score: 1830.50 Matches: 355
Percent Similarity: 69.87% Conservative: 74
Best Local Similarity: 57.82% Mismatches: 154
Query Match: 56.31% Indels: 31
DB: 4 Gaps: 10

US-09-807-802A-15 (1-599) x US-09-532-594B-1 (1-4767)

Qy 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20
Db 2668 ACGGCTCCTGGAAAGAGAGACCGCTTGATTGAATCCCCCAGCAGCCCGACTCCTCCACG 2727
Qy 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPhe---GlyGlnThr 39
Db 2728 GGTATCGGCAAAAAGGCAAGCAGCCGCTAAAAAGAGCTCGTTTTCGAAGACGAAACT 2787
Qy 40 GlyAspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAla 59
Db 2788 GGAGCAGCGCAGCGACCCCTGAGGGATCAACTCCGGA-----GCC 2829
Qy 60 AlaValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGlu 79
Db 2830 ATGCTGTATGACAGTGAGATGCGTGACAGCAGCTGGCGGAGCTGCAGTCGAGGGGGACAA 2889
Qy 80 GlyAlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGly 99
Db 2890 GGTGCCGATGGAGTGGGTAAATGCCTCGGGTGATGGCATTGGCATTCCACCTGGTCTGAG 2949
Qy 100 AspArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeu 119
Db 2950 GGCCACGTCACGACCACAGCAGCAGCAGACCTGGGTCTTGCCACCTACAACACCACCTN 3009
Qy 120 TyrLysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyr 139
Db 3010 TACAAGCGACTCGGAGAG-----AGCCTGCAGTCCAAACACCTACAACCGGATTTC 3057
Qy 140 SerThrProTyrGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAsp 159
Db 3058 TCCACCCCTGGGGATACTTTGACTTCAACCGCTTCCACTGCCACTTCTCACCACGTGAC 3117
Qy 160 TrpGlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeu 179
Db 3118 TGGCAGCGACTCATCAACAACAACACTGGGGCATCGGACCCCAAGCCATCGCGGTCAAATC 3177
Qy 180 PheAsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsn 199
Db 3178 TTCAACATCCAGGTCAAGAGGTACGACGTCGAAAGCGGAGACAAACGGTGGCTAATAAC 3237
```

```
Qy 200 LeuThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGly 219
Db 3238 CTTACCAGCACGGTTCAGATCTTTGCGGACTCGTCGTACGAACCTGCCGTACGTGATGGAT 3297
Qy 220 SerAlaHisGlnGlyCysLeuProPheProPheProAlaAspValPheMetIleProGlnTyr 239
Db 3298 GCGGGTCAAGAGGGCAGCCTGCCTCTTTTCCCAACGACGCTCTTTATGGTGGCCAGTAC 3357
Qy 240 GlyTyr-----LeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPhe 256
Db 3358 GGCTACTGTGGACTGGTGACCGGCAACACTTCGCGACCAACAGACTGACAGAAATGCCTTC 3417
Qy 257 TyrCysLeuGluTyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSer 276
Db 3418 TACTGCCTGGAGTACTTCTTCGAGATGCTCGGAGTGGCAACACTTTGAAATTACG 3477
Qy 277 TyrThrPheGluGluValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArg 296
Db 3478 TACAGTTTGGAGAGGTGCCTTTCCACTCGATGACGGGCACAGCCAGAGCCTGGACCGG 3537
Qy 297 LeuMetAsnProLeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsn----- 314
Db 3538 CTGATGAACCTCTCATCGACCAGTACCTGTGGGACTGCAATCGACCACCCACCGGAACC 3597
Qy 315 -----GlnSerGlySerAlaGlnAsnLysAspLeuLeuPheSerArgGlySerProAla 332
Db 3598 ACCCTGAATGCCGGGACTGCCACCACCAAC-----TTTACCAAGCTGCGGCTTACC 3648
Qy 333 GlyMetSerValGlnProLysAsnTrpLeuProGlyProCysTyrArgGlnGlnArgVal 352
Db 3649 AACTTTTCCAACCTTTAAAAGAACTGGCTGCCCGGGCTTCAATCAAGCAGCAGGGCTTC 3708
Qy 353 SerLysThrLysThrAspAsnAsnAsnSerAsnPheThrTrpThrGlyAlaSer----- 370
Db 3709 TCAAAAGACTGCC-----AATCAAAACTACAAGATCCTGCCACCGGGTCAGACAGTCTC 3762
Qy 371 ---LysTyr-----AsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 385
Db 3763 ATCAAAATACGAGACGACAGCAGTCTTGACGGAAGATGGAGTGCCTGACCCCGGACCT 3822
Qy 386 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 405
Db 3823 CCAATGCCACGGCTGGACTGCGGACAGCAAG--TTTCAGCAACAGCCAGCTCATCTTT 3879
Qy 406 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 425
Db 3880 GCGGGCCTAAACAGAACGCAACACGCGCCACCGTACCCGGGACTCTGATCTTCACCTCT 3939
Qy 426 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 445
Db 3940 GAGGAGAGCTGGCAGCCACCAACGCCACCGATACGGACATGTGGGGCAACCTACCTGSC 3999
Qy 446 AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 465
Db 4000 GGTGACAGAGCAACAGCAACCTGCCGACCGTGGACAGACTGACAGCCTTGGGAGCCGTG 4059
Qy 466 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 485
Db 4060 CCTGGAATGGTCTGGCAAAACAGAGACATTTACTACCGGTGCCATTTGGGCTGAAACAC 4119
Qy 486 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 505
Db 4120 CCTCATACCGATGGACACTTTTACCCCTCACCCTGATTTGGTGGTGGTGGCTGAAACAC 4179
Qy 506 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 525
Db 4180 CCGCCTCTCTCAATTTTATCAAGAACACCCCGGTACCTGCGAATCCTGCAACGACCTTC 4239
Qy 526 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 545
Db 4240 AGCTCTACTCCGGTAAACTCCTTCATTACTCAGTACAGCAGTGGCCAGGTGTGCGTGCAG 4299
Qy 546 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 565
```

Db 4300 ATTGACTGGGAGATCCAGAAGGAGCGGTCCAAACGCTGGAACCCCGAGGTCAGATTACC 4359  
Qy 566 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 585  
Db 4360 TCCAACTACGGACAGCAAAACTCTCTGTGTGGGCTCCCGATCGCGCTGGGAAATACACT 4419  
Qy 586 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 599  
Db 4420 GAGCCTAGGGCTATCGTACCGCTACCTCACCCACCACCTG 4461

RESULT 13  
US-09-438-268-1/c  
; Sequence 1, Application US/09438268  
; Patent No. 6491907  
; GENERAL INFORMATION:  
; APPLICANT: Rabinowitz, Joseph E.  
; APPLICANT: Samulski, Richard J  
; APPLICANT: Xiao, Weidong  
; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING  
; TITLE OF INVENTION: THE SAME  
; FILE REFERENCE: 5470-186  
; CURRENT APPLICATION NUMBER: US/09/438,268  
; CURRENT FILING DATE: 1999-11-10  
; EARLIER APPLICATION NUMBER: 60/107,840  
; EARLIER FILING DATE: 1998-11-10  
; EARLIER APPLICATION NUMBER: 60/123,651  
; EARLIER FILING DATE: 1999-03-10  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 7214  
; TYPE: DNA  
; ORGANISM: Virus  
US-09-438-268-1

Alignment Scores:  
Pred. No.: 9.21e-167 Length: 7214  
Score: 1830.50 Matches: 355  
Percent Similarity: 69.87% Conservative: 74  
Best Local Similarity: 57.82% Mismatches: 154  
Query Match: 56.31% Indels: 31  
DB: 4 Gaps: 10  
  
US-09-807-802a-15 (1-599) x US-09-438-268-1 (1-7214)  
  
Qy 1 ThrAlaProGlyLysLysArgProValGluGlnSerProGlnGluProAspSerSerSer 20  
Db 2575 ACGGCTCCTGGAAAGAGAGACCGTTGATTGAATCCCCCGCAGCAGCCGACTCCTCCACG 2516  
Qy 21 GlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPhe---GlyGlnThr 39  
Db 2515 GGTATCGGCAAAAGGCAAGCAGCCGGCTAAAGAGAGCTCGTTTCGAAGACGAAACT 2456  
Qy 40 GlyAspSerGluSerValProAspProGlnProLeuGlyGluProProAlaThrProAla 59  
Db 2455 GGAGCAGCGACGGACCCCTGAGGGATCAACTTCCGGA-----GCC 2414  
Qy 60 AlaValGlyProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGlu 79  
Db 2413 ATGTCTGATGACAGTGAGATGCGTGACAGCAGCTGGCGGAGCTGCAGTCGAGGGCGGACAA 2354  
Qy 80 GlyAlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGly 99  
Db 2353 GGTGCCGATGGAGTGGGTAATGCTCGGGTGATTGGCATTGCGATTCCACCTGGTCTGAG 2294  
Qy 100 AspArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeu 119  
Db 2293 GGCCACGTACGACCACCGACAGCAACCTGGGTCTTGCCACCTTACCAACACCACTC 2234  
Qy 120 TyrLysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyr 139  
Db 2233 TACAGCGACTCGGAGAG-----AGCCTGCAGTCCAAACACCTACACCGGATTC 2186

Qy 140 SerThrProTrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAsp 159  
Db 2185 TCCACCCCTCGGGATACTTTGACTTCAACCGCTTCCACTGCCACTTCTCACCACGTGAC 2126  
Qy 160 TrpGlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeu 179  
Db 2125 TGGCAGCGACTCATCAACAACAACTGGGCGCATGCGACCCAAAGCCATGCGGGTCAAAATC 2066  
Qy 180 PheAsnIleGlnValLysGluValThrThrAsnAspGlyValThrThrIleAlaAsnAsn 199  
Db 2065 TTCAACATCCAGGTCAAGGAGGTACAGACTCGAAGCTCGAAGCTGCGTAAATAC 2006  
Qy 200 LeuThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGly 219  
Db 2005 CTTACCAGCACGGTTCAGATCTTTGCGGACTCGTACGAACTGCGGTAGTGATGAT 1946  
Qy 220 SerAlaHisGlnGlyCysLeuProProPheProAlaAspValPheMetIleProGlnTyr 239  
Db 1945 GCGGGTCAAGAGGGCAGCCTGCCCTCTTTTCCCAACGACGTCTTTATGSGTCCCCAGTAC 1886  
Qy 240 GlyTyr-----LeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPhe 256  
Db 1885 GGCTACTGTGGACTGGTGACCGGCAACACTTCCGAGCAACAGACTGACAGAAATGCCTTC 1826  
Qy 257 TyrCysLeuGluTyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSer 276  
Db 1825 TACTGCCTGGAGTACTTTCCTTCGACAGATGCTGCGGACTGGCAACAACATTGAAATTACG 1766  
Qy 277 TyrThrPheGluGluValPropheHisSerSerTyrAlaHisSerGlnSerLeuAspArg 296  
Db 1765 TACAGTTTGTAGAAGGTGCCTTTCCACTCGATGTACGCGCACAGCCAGAGCCTGGACCGG 1706  
Qy 297 LeuMetAsnProLeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsn----- 314  
Db 1705 CTGATGAACCTCTCATCGACCACTGCTGCGGACTGCAATCGACCAACACCGGAACC 1646  
Qy 315 -----GlnSerGlySerAlaGlnAsnLysAspLeuLeuPheSerArgGlySerProAla 332  
Db 1645 ACCCTGAATGCGGGACTGCCACCACCAAC-----TTTACCAAGCTGCGGCTTACC 1595  
Qy 333 GlyMetSerValGlnProLysAsnTrpLeuProGlyProCysTyrArgGlnGlnArgVal 352  
Db 1594 AACTTTTCCAACCTTAAAAAGAACTGGCTGCGCGGCTTCAATCAAGCAGCAGGGCTTC 1535  
Qy 353 SerLysThrLysThrAspAsnAsnAsnSerAsnPheThrTrpThrGlyAlaSer----- 370  
Db 1534 TCAAAGACTGCC-----AATCAAAACTACAAGATCCCTGCCACCGGTCAGACAGTCTC 1481  
Qy 371 ---LysTyr-----AsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 385  
Db 1480 ATCAAATACGAGACGACAGCACTCTGGACGGAAGATGGAGTGCCTGACCCCGGACCT 1421  
Qy 386 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 405  
Db 1420 CCAATGGCCACGGCTGGACCTGCGGACAGCAAG---TTCAGCAACAGCCAGCTCATCTTT 1364  
Qy 406 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 425  
Db 1363 GCGGGGCTTAAACAGAACGGCAACACGCGCCACCGTACCGGACTCTGATCTTCACTCT 1304  
Qy 426 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 445  
Db 1303 GAGGAGGAGCTGGCAGCCACCAACGCCACCGATACGGACATGTGGGGCAACCTACCTGGC 1244  
Qy 446 AsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 465  
Db 1243 GGTGACCAGAGCAACAGCAACCTGCCGACCCCTGGACAGACTGACAGCCTTGGGAGCCGTG 1184  
Qy 466 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 485  
Db 1183 CCTGGAATGGTCTGGCAAAACAGAGACATTTACTACAGGGGTCCCATTTTGGGCCAAGATT 1124



|    |      |  |      |
|----|------|--|------|
| Qy | 486  | ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn   | 505  |
| Db | 1123 | CCTCATACCGATGGACACTTTCACCCCTCACCCGCTGATTGGTGGTTTGGCTGAAACAC    | 1064 |
| Qy | 506  | ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe   | 525  |
| Db | 1063 | CCGCCCTCTCAAAATTTTATCAAGAACAACCCCGGTACCTGCGAATCTCTGCAACGACCTTC | 1004 |
| Qy | 526  | SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu   | 545  |
| Db | 1003 | AGCTCTACTCCGGTAAACTCCTTCATTACTCAGTACAGCACTGCCAGGTGTCGGTGCGAG   | 944  |
| Qy | 546  | IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr   | 565  |
| Db | 943  | ATTGACTGGGAGATCCAGAAGGAGCGGTCCAAACGCTGGAACCCCGAGGTCCAGTTTACC   | 884  |
| Qy | 566  | SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr   | 585  |
| Db | 883  | TCCAACTACGGACAGCAAAACTCTCTGTGTGGGCTCCCGATCGGGCTGGGAAATACACT    | 824  |
| Qy | 586  | GluProArgProIleGlyThrArgTyrLeuThrArgProLeu                     | 599  |
| Db | 823  | GAGCCTAGGGCTATCGGTACCCGCTACCTCACCCACCACCTG                     | 782  |

## RESIN, T 14

```

US-09-532-594B-19
; Sequence 19, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.025202
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 1617
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: synthetic construct
; NAME/KEY: misc_feature
; LOCATION: 162
; OTHER INFORMATION: n = g, a, c or t(u)
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 capsid protein vp3 gene

```

```

Alignment Scores:
Pred. NO.:      3.19e-154
Score:          1690.50
Percent Similarity: 71.30%
Best Local Similarity: 59.41%
Query Match:    52.00%
DB:             4

US-09-807-802A-15 (1-599) x US-09-532-594B-19 (1-1617)

Qy   67 AlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGlyAsn 86
     |||:::|||||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db   7 GCAGCAGCTGGCGGAGCTGCAGTC--GAGGSGGACAAAGGTGCCGATGGAGTGGGTAAT 63

Qy   87 AlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThrSer 106
     |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db   64 GCCTCGGGTGATTGGCATTTCCGATTCCACCTGGTCTGAGGGGCCACGTCACGACCACCAGC 123

Qy   107 ThrArgThrTrpAlaLeuProThrTyraAsnAsnHisLeuTyrlLysGlnIleSerSerAla 126
       |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

|    |      |  |      |
|----|------|--|------|
| Db | 124  | ACCAGAACCTGGGTCTTGGCCACCTACAAACAACCACCTNTACAAAGCGACTCGGAGAG--- | 180  |
| QY | 127  | SerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTyrGlyTyrPhe   | 146  |
| Db | 181  | -----AGCCTGCAGTCCAAACACCTACAAACGGATTCTCCACCCCTGGGATACTTT       | 231  |
| QY | 147  | AspPheAsnArgPheHisCysHisPheSerProArgAspTyrGlnArgLeuIleAsnAsn   | 166  |
| Db | 232  | GACTTCAACCGCTTCCACTGCCACTTCTCACCAACGCTGACTGGCAGCACTCATCAACAAC  | 291  |
| QY | 167  | AsnTyrGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLysGlu   | 186  |
| Db | 292  | AACCTGGGCGATCGACCCCAAGCCATCGGGTCAAAATCTTCAACATCCAGGTCAAGGAG    | 351  |
| QY | 187  | ValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGlnVal   | 206  |
| Db | 352  | GTCACGACGTGCAACGGCGAGACAACGGTGGCTAATAACCTTACCAGCACGGTTTCAGATC  | 411  |
| QY | 207  | PheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCysLeu   | 226  |
| Db | 412  | TTTGGGACTCGTCTGACGAACCTGCGTACGTGATGGATGCGGGTCAAGAGGGCAGCCTG    | 471  |
| QY | 227  | ProProPheProAlaAspValPheMetIleProGlnTyrGlyTyr-----LeuThr       | 243  |
| Db | 472  | CCTCCTTTTCCCAACGACGTCTTTATGGTGTGCCCCAGTACGGCTACTGTGGACTGGTGACC | 531  |
| QY | 244  | LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro   | 263  |
| Db | 532  | GGCAACACTTCGACGACCAACAGACTGCACAGAAATGCCTTCTACTGCCTGGAGTACTTTCT | 591  |
| QY | 264  | SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro   | 283  |
| Db | 592  | TCGCAGATGCTGCGGACTGGCAACAACCTTGAAATTACGTACAGTTTGGAGAAGGTGCCT   | 651  |
| QY | 284  | PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp   | 303  |
| Db | 652  | TTCCACTCGATGTACGGCGACAGCCAGAGCGCTGGACCGCTGATGAACCTCTCATCGAC    | 711  |
| QY | 304  | GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsn-----GlnSerGlySerAla          | 319  |
| Db | 712  | CAGTACCTGTGGGACTGCAATCGACCACCCCGGAACCCCTGAATGCCGGGACTGCC       | 771  |
| QY | 320  | GlnAsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLys   | 339  |
| Db | 772  | ACCACCAAC-----TTTACCAAGCTGCGGCCCTACCAACTTTTCCAACTTTAAAAAG      | 822  |
| QY | 340  | AsnTyrLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsn   | 359  |
| Db | 823  | AACCTGGCTGCCCGGCCTTCAATCAAGCAGCAGGGCTTCTCAAGAGACTGCC-----AAT   | 876  |
| QY | 360  | AsnAsnSerAsnPheThrTyrThrGlyAlaSer-----LysTyr-----              | 372  |
| Db | 877  | CAAAACTACAAGATCCCTGCCACCGGGTCAGACAGTCTCATCAAAATACGAGACGCACAGC  | 936  |
| QY | 373  | AsnLeuAsnGlyArgGluSerIleIleAsnProGlyThrAlaMetAlaSerHisLysAsp   | 392  |
| Db | 937  | ACTCTGGACGGAGATGGAGTGCCCTGACCCCGGACCTCCAATGGCCACGGGTGGACCT     | 996  |
| QY | 393  | AspGluAspLysPhePheProMetSerGlyValMetIlePheGlyLysGluSerAlaGly   | 412  |
| Db | 997  | GCGGACAGCAAG---TTCAGCAACAGCCAGCTCATCTTTTGGGGGCCCTAAACAGAACGGC  | 1053 |
| QY | 413  | AlaSerAsnThrAlaLeuAspAsnValMetIleThrAspGluGluIleLysAlaThr      | 432  |
| Db | 1054 | AACACGGCCACCGTACCGGGACTCTGATCTTACCTCTGTAGGAGGAGCTGGCAGCCACC    | 1113 |
| QY | 433  | AsnProValAlaThrGluArgPheGlyThrValAlaValAsnPheGlnSerSerThr      | 452  |
| Db | 1114 | AACGCCACCGATACGGACATGTGGGGCAACCTTACCTGGCGGTGACCAGAGCAACAGCAAC  | 1173 |
| QY | 453  | AspProAlaThrGlyAspValHisAlaMetGlyAlaLeuProGlyMetValTyrGlnAsp   | 472  |
| Db | 1174 | CTGCCGACCGTGGACAGACTGACAGCCTTGGGAGCGCTGCCTGGAATGGTCTGGCAAAAC   | 1233 |

|    |      |  |      |
|----|------|--|------|
| QY | 473  | ArgAspValTyrLeuGlnGlyProIleTirAlaLysIleProHisThrAspGlyHisPhe   | 492  |
| DB | 1234 | AGAGACATTACTACAGGGTCCCATTTGGGCCAAGATTCTCATACCGATGGACACTTT      | 1293 |
| QY | 493  | HisProSerProLeuMetGlyGlyPheGlyLeuLysAsnProProGlnIleLeuIle      | 512  |
| DB | 1294 | CACCCCTCACCGCTGATTGGTGGTTTGGGCTGAACACCCGCCCTCCTCAAATTTTATC     | 1353 |
| QY | 513  | LysAsnThrProValProAlaAsnProProAlaGluPheSerAlaThrLysPheAlaSer   | 532  |
| DB | 1354 | AAGAACACCCCGGTACCTGCGAATCTCTGCAACGACCTTCAGCTCTACTCCGGTAAACTCC  | 1413 |
| QY | 533  | PheIleThrGlnTyrSerThrGlyGlnValSerValGluIleGluTirPLeuLeuGlnLys  | 552  |
| DB | 1414 | TTCATTACTCAGTACAGCACTGGCCAGGTGTCGGTGCAGATTGACTGGGAGATCCAGAAG   | 1473 |
| QY | 553  | GluAsnSerLysArgTirAsnProGluValGlnTyrThrSerAsnTyrAlaLysSerAla   | 572  |
| DB | 1474 | GAGCGGTCCAAACGCTGGAAACCCCGAGGTCCAGTTTACCTCCAACCTACGGACAGCAAAAC | 1533 |
| QY | 573  | AsnValAspPheThrValAspAsnAsnGlyLeuTyrThrGluProArgProIleGlyThr   | 592  |
| DB | 1534 | TCTCTGTTGTGGGCTCCCGATCGCGTGGGAAATACACTGAGCCTAGGGCTATCGGTACC    | 1593 |
| QY | 593  | ArgTyrLeuThrArgProLeu  | 599  |
| DB | 1594 | CGCTACCTCACCCACCACCTG  | 1614 |

**RESULT 15**

US-09-438-268-3  
; Sequence 3, Application US/09438268  
; Patent No. 6491907  
; GENERAL INFORMATION:  
; APPLICANT: Rabinowitz, Joseph E.  
; APPLICANT: Samulski, Richard J  
; APPLICANT: Xiao, Weidong  
; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING  
; TITLE OF INVENTION: THE SAME  
; FILE REFERENCE: 5470-186  
; CURRENT APPLICATION NUMBER: US/09/438,268  
; CURRENT FILING DATE: 1999-11-10  
; EARLIER APPLICATION NUMBER: 60/107,840  
; EARLIER FILING DATE: 1998-11-10  
; EARLIER APPLICATION NUMBER: 60/123,651  
; EARLIER FILING DATE: 1999-03-10  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 2271  
; TYPE: DNA  
; ORGANISM: Virus  
US-09-438-268-3

Db 1582 ATATCCTTAGCCCTGGGCCAGTGTCTCAGCCATACCACCCTGGGACACAGATAAATAT 1641  
Qy 398 PheProMetSerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAla 417  
Db 1642 GTCACAGGAATAAATGCCATTCTCATGGTCAGACCACCTTATGGT-----AACGCT 1692  
Qy 418 LeuAspAsnVal-----MetIleThrAspGluGluGluIleLys 430  
Db 1693 GAAGACAAAGAGTATCAGCAAGGAGTGGGTAGATTTCCAAATGAAAAAGACAGCTAAAA 1752  
Qy 431 AlaThrAsnProValAlaThrGluArgPheGlyThrValAlaValAsnPheGlnSerSer 450  
Db 1753 CAGTTACAGGGTTTAAACATGCACACCTAC-----TTTCCCAATAAA 1794  
Qy 451 SerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeuProGlyMetValTrp 470  
Db 1795 GGAACCCAGCAATATACAGATCAAAATTGAG---CGCCCCCTAATGGTGGTTCTGTATGG 1851  
Qy 471 GlnAspArgAspValTyrLeuGlnGlyProIleTyrAlaLysIleProHisThrAspGly 490  
Db 1852 AACAGAGAGCCCTTCACTATGAAGCCAGCTGTGGAGTAAATTCCAAATTTAGATGAC 1911  
Qy 491 HisPheHisProSer---ProLeuMetGlyGlyPheGlyLeuLysAsnProProProGln 509  
Db 1912 AGTTTTTAAACTCAGTTTGCAGCCCTTAGGAGGATGGGGTTTGCATCAGCCACCTCCTCAA 1971  
Qy 510 IleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPheSerAlaThrLys 529  
Db 1972 ATATTTTAAAAA-----ATATTACCAAAAGTGGCCCAATTGGAGGTATTAAATCAATG 2025  
Qy 530 PheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGluIleGluTrpGlu 549  
Db 2026 GGAATTACTACTAGTTTCACTAGTATGCCGTGGGAATTATGACAGTAACATGACATTTAAA 2085  
Qy 550 Leu---GlnLysGluAsnSerLysArgTrpAsnProGlu 561  
Db 2086 TTGGGGCCCCCGTAAAGCTACGGGACGGTGGAAATCCTCAA 2124

Search completed: February 15, 2004, 01:37:39  
Job time : 249.456 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 14, 2004, 15:26:50 ; Search time 3324.15 Seconds  
(without alignments)  
4379.586 Million cell updates/sec

Title: US-09-807-802A-15  
Perfect score: 3251  
Sequence: 1 TAPGKKRPVEQSPQEPDSSS.....NNGLYTEPRPICTRYLTRPL 599

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09807802/runat\_11022004\_175608\_15941/app\_query.fasta\_1.2389  
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09807802@cgn\_1\_17257@runat\_11022004\_175608\_15941 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_ges1:\*

29: gb\_ges2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID       | Description         |
|------------|-------|-------|--------|----|----------|---------------------|
| 1          | 140.5 | 4.3   | 2318   | 11 | BC014681 | BC014681 Homo sapi  |
| 2          | 134   | 4.1   | 856    | 28 | BH164736 | BH164736 ENTT28TP   |
| 3          | 125.5 | 3.9   | 2146   | 11 | AK042727 | AK042727 Mus muscu  |
| 4          | 125.5 | 3.9   | 2295   | 11 | AK035953 | AK035953 Mus muscu  |
| 5          | 125   | 3.8   | 754    | 10 | BG207226 | BG207226 RST26694   |
| 6          | 121.5 | 3.7   | 925    | 28 | BH152154 | BH152154 ENTPV12TR  |
| 7          | 115.5 | 3.6   | 753    | 28 | BH115587 | BH115587 RPCI-24-3  |
| 8          | 114.5 | 3.5   | 1042   | 13 | BQ642977 | BQ642977 AGENCOURT  |
| 9          | 114   | 3.5   | 5809   | 11 | AK076994 | AK076994 Mus muscu  |
| 10         | 113.5 | 3.5   | 2637   | 11 | BC037597 | BC037597 Mus muscu  |
| 11         | 113.5 | 3.5   | 2688   | 11 | BC019128 | BC019128 Mus muscu  |
| 12         | 113   | 3.5   | 1960   | 11 | AY103595 | AY103595 Zea mays   |
| 13         | 113   | 3.5   | 4782   | 11 | AK048546 | AK048546 Mus muscu  |
| 14         | 112.5 | 3.5   | 1651   | 11 | AY104980 | AY104980 Zea mays   |
| 15         | 112   | 3.4   | 3057   | 11 | AK044947 | AK044947 Mus muscu  |
| 16         | 111.5 | 3.4   | 3329   | 11 | AK033012 | AK033012 Mus muscu  |
| 17         | 111.5 | 3.4   | 4124   | 11 | AK031710 | AK031710 Mus muscu  |
| 18         | 110.5 | 3.4   | 768    | 14 | CD558053 | CD558053 AGENCOURT  |
| 19         | 110.5 | 3.4   | 815    | 10 | BG746848 | BG746848 602704187  |
| 20         | 110.5 | 3.4   | 925    | 13 | BQ440075 | BQ440075 AGENCOURT  |
| 21         | 110.5 | 3.4   | 2904   | 11 | AK046602 | AK046602 Mus muscu  |
| 22         | 110.5 | 3.4   | 3619   | 11 | AK028748 | AK028748 Mus muscu  |
| 23         | 110.5 | 3.4   | 4843   | 11 | AK082447 | AK082447 Mus muscu  |
| 24         | 110   | 3.4   | 640    | 14 | CA356072 | CA356072 628112 NC  |
| 25         | 110   | 3.4   | 701    | 14 | CA363789 | CA363789 638563 NC  |
| 26         | 110   | 3.4   | 703    | 14 | CA378088 | CA378088 656813 NC  |
| 27         | 110   | 3.4   | 902    | 29 | CNS03388 | AL225809 Tetraodon  |
| 28         | 109.5 | 3.4   | 2365   | 11 | AK081012 | AK081012 Mus muscu  |
| 29         | 109   | 3.4   | 350    | 9  | AW545004 | AW545004 C0188E06 - |
| 30         | 109   | 3.4   | 581    | 13 | BQ163384 | BQ163384 952077G03  |
| 31         | 109   | 3.4   | 659    | 13 | BW314066 | BW314066 BW314066   |
| 32         | 109   | 3.4   | 687    | 13 | BU723756 | BU723756 SJMAZH04   |
| 33         | 108.5 | 3.3   | 838    | 10 | BE642450 | BE642450 Cri2 6 B0  |
| 34         | 108   | 3.3   | 497    | 12 | BM417785 | BM417785 952005F11  |
| 35         | 108   | 3.3   | 2125   | 11 | AK079113 | AK079113 Mus muscu  |
| 36         | 108   | 3.3   | 2931   | 11 | AK019448 | AK019448 Mus muscu  |
| 37         | 108   | 3.3   | 4784   | 11 | AK041115 | AK041115 Mus muscu  |
| 38         | 107   | 3.3   | 411    | 12 | BM152500 | BM152500 TCBAP1E91  |
| 39         | 107   | 3.3   | 544    | 12 | BM896109 | BM896109 952063E01  |
| 40         | 107   | 3.3   | 550    | 29 | CC171347 | CC171347 i171d11.b  |
| 41         | 107   | 3.3   | 576    | 12 | BM173665 | BM173665 952003C01  |
| 42         | 107   | 3.3   | 582    | 12 | BM417581 | BM417581 952018G12  |
| 43         | 107   | 3.3   | 582    | 13 | BU049515 | BU049515 1111009D0  |
| 44         | 107   | 3.3   | 588    | 12 | BM895805 | BM895805 952065E09  |
| 45         | 107   | 3.3   | 672    | 12 | BM333577 | BM333577 MEST157-G  |

ALIGNMENTS

RESULT 1  
BC014681  
LOCUS BC014681  
DEFINITION Homo sapiens, Similar to osa, clone IMAGE:3866722, mRNA.  
ACCESSION BC014681  
VERSION BC014681.1 GI:15928395  
KEYWORDS HTC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2318)  
AUTHORS Strausberg, R.

TITLE Direct Submission  
JOURNAL Submitted (01-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [villalon@bcm.tmc.edu](mailto:villalon@bcm.tmc.edu)  
Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 21 Row: 1 Column: 9  
This clone was selected for full length sequencing because it passed the following selection criteria: GenomesCan gene prediction  
This clone has the following problem: frame shifted.

FEATURES

Location/Qualifiers  
source 1..2318  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3866722"  
/tissue type="Eye, retinoblastoma"  
/clone\_lib="NIH\_MGC\_67"  
/lab\_host="DH10B"  
/note="vector: pCMV-SPORT6"

BASE COUNT 637 a 680 c 544 g 457 t

ORIGIN

Alignment Scores:

Pred. No.: 0.0831 Length: 2318  
Score: 140.50 Matches: 143  
Percent Similarity: 33.38% Conservative: 77  
Best Local Similarity: 21.70% Mismatches: 213  
Query Match: 4.32% Indels: 228  
DB: 11 Gaps: 36

US-09-807-802A-15 (1-599) x BC014681 (1-2318)

QY 13 ProGlnGluProAspSerSerGlyIleGlyLysThrGlyGlnGlnProAlaLysLys 32  
Db 718 CCTCAGGTGCCCCATGGTGGCAGTGGTGGCGGTTCAGATGGGTGTCTACCCCTGGCATGCAG 777  
QY 33 ArgLeuAsnPheGlyGlnThr---GlyAspSerGluSer----- 44  
Db 778 AATCAGAGGCATGGGCAATCCTTTGTGGACAGCAGCTCCATGTGGGGCCCCAGGGCTGTT 837  
QY 45 ---ValProAsp-----ProGlnProLeuGlyGluProPro 55  
Db 838 CAGGTACCAGACACAGATACGAGCCCCCTTACCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 897  
QY 56 AlaThrProAlaAlaValGlyPro----- 63  
Db 898 ---CAGCCGGCTCCGTCGGGGCCCCCTGTACAGGGCCACCTCAGCACATGCAGCAGATG 954  
QY 64 ---ThrThrMetAlaSerGlyGly-----GlyAlaPro----- 73  
Db 955 GGCAGCTATATGGCAGCGTGGGGATTTTTCATGCAGCAGCATGGTCAGCCACAGCAGCAGG 1014  
QY 74 MetAlaAspAsnAsnGluGlyAlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCys 93  
Db 1015 ATGAGCCAGTTTTCCTCCAGAGCCACAGAGGGCCTCAATCAG----- 1053

QY 94 AspSerThrTrpLeuGlyAspArgValIleThrThrSerThrArgThrTrpAlaLeuPro 113  
Db 1054 -----GGAAATCCTTTTATTGGCCACCTCAGG-----CCT 1083  
QY 114 ThrTyrAsnAsnHisLeuTyrLysGlnIleSerSerAlaSerThrGlyAlaSerAsnAsp 133  
Db 1084 GGCCACTTGTCCACGTCGCCACGACAGAGTCCCAGCATGGCACCTTCTCTGGTCACTCG 1143  
QY 134 -----AsnHisTyrPheGlyTyrSerThrProTrpGlyTyrPheAspPheAsn 149  
Db 1144 GTGCAGCAGTTCCATCACCAACC-CTCTACTGTCTCTCCATGG----- 1184  
QY 150 ArgPheHisCys-His-----PheSerProArgAspTrpGlnArgLeuIleAsnAs 166  
Db 1185 AGAATCCGTTGCCACAGTCCCAGATTCTCCCCGAATCTCCCCCAACAA----- 1233  
QY 166 nAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLysG 186  
Db 1234 ---GGGGCTGTTAGGCCGCAACCTTAACCTTCTCGGAGCCAGACAGTCCCTC 1289  
QY 186 uValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGlnVa 206  
Db 1290 TCCTACTATAACAAC----- 1305  
QY 206 lPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCysLe 226  
Db 1306 ---TCAGGGCAGTATTCTCGATATCCTTAC-----AGTAACCTAATCAGGGATTAGT 1355  
QY 226 u-----ProProPheProAlaAs 232  
Db 1356 TAACAATACAGGGATGAATCAAAATTTAGGCCCTTACAAATAATACTCCAATGAATCAGTC 1415  
QY 232 pValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsnAsnGlySerGlnAlaValG 252  
Db 1416 CGTA-----CCAAGATAC-----CCCAATGCTGTAGG 1442  
QY 252 yArgSerSerPheTyrCysLeuGluTyrPheProSerGlnMetLeuArgThrGlyAsnAs 272  
Db 1443 A-----TTCCCATCA-----AACAGTGGTCAAGG 1466  
QY 272 nPheThrPheSerTyrThrPheGluGluValProPheHisSerSerTyrAlaHisSerG 292  
Db 1467 ACTA-----ATGCACCAGCAGCCCCATCCACCCAGTGGCTCACTTAACCA 1511  
QY 292 nSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyrLeuTyrTyrLeuAsnArgTh 312  
Db 1512 AATGAACACACAAACTATGCATCCT-----TCACAGCC 1544  
QY 312 rGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeuPheSerArgGlySerProAl 332  
Db 1545 TCAGGGAACCTTATGCCTCTCCACCTCCCATGTCCACCCATGAAGCAATGAGTAATCCAGC 1604  
QY 332 aGlyMetSerValGlnProLysAsnTrpLeuProGlyProCysTyrArgGlnGlnArgVa 352  
Db 1605 AGGC-----ACTCCTCTCTCCACAAGTCAGCCCGGAAG-TGCTGG----- 1643  
QY 352 lSerLysThrLysThrAspAsnAsnAsnSerAsnPheThrTrpThrGlyAlaSerLysTy 372  
Db 1644 -----GATACCAATGGAAGTTGGCAG---TTATCCAAATA 1675  
QY 372 rAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThrAlaMetAlaSerHisLysAs 392  
Db 1676 T-----GCCCATCCTCAGCCATCTCACCAGCCCCCTGGTGGCATGGGAATCGGAC 1726  
QY 392 pAspGluAspLysPhePheProMetSerGlyValMetIlePheGlyLysGluSerAlaG 412  
Db 1727 AGAGGAA-----TATGGGCCCCCAGAAACATGC 1753  
QY 412 yAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp-----GluGluG 428  
Db 1754 AGCAGTCTCGTCCATTTATAGGCATGTCCTCGGCACCAAGGAATTGACTGGGCACATGA 1813  
QY 428 uIleLys-----AlaThrAsnProValAlaThrG 438

```

|  |||
1814 GGCCAAATGGTTGCTCCTGGTGTGGCCCTGGAGACCCCAAGCAATCCAGGAACGACTGA 1873
QY  uArgPheGlyThrValAlaValAlaAsnPheGlnSerSerSerThrAspProAlaThrGlyAs 458
Db  :  ::  |||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
1874 TACCTGGCCCAACAACATCC-----TGGTCAACAGCCATCTTTTCAGCAGTTGCCAA 1924
QY  pValHisAlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuG1 478
Db  :  ::  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1925 CTGTCTCTCCACTGCAGCCTCACCCGGG-----CTTGC 1957
QY  nGlyProIleTrpAlaLysIle-ProHisThrAsp-----GlyHisPheHisProSerP 496
Db  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
1958 ACCACCAGTCTTCCACCTCCACACCCCTCATCCAGCCTTGGGCACAGCTCCACCCATCAC 2017
QY  robeuMetGlyGlyPheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrP 516
Db  ||  ||  ::  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
2018 CC-----CAGAACACCCCGCAGAAAGTG-----C 2041
QY  roValProAlaAsnProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrG 536
Db  ||||  |||  ::  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
2042 CTGTGCATCAGCATTCCTCCCGTCGGAGCCCTTCTTAGAGAAACAGTGCCCGGATGACTC 2101
QY  536 lnTyrSerThrGlyGlnValSerVal-----GluIleGluT 548
Db  ||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
2102 AGGTAGTGGACCGAATGCTCAGCTAGTGAAGAGTGATGATTACCTGCCATCAATAGAAC 2161
QY  548 rpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThrSerAsnT 568
Db  ::  |||  ::  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
2162 AGCAGCCACAACAAAGAGAGAGAAAGAAACCAACCATTTGTAGCAGAGATCCCA 2221
QY  568 yzAlaLysSerAlaAsnValAspPheThr-----ValAspAsnAsnGlyLeu 583
Db  ||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
2222 GTAAAGGTTTGTGTAAGATGACTTCCTCGTGGGTAGATAACCAAGAACTA 2274
```

```

RESULT 2
BH164736
LOCUS      ENTTS28TF Entamoeba histolytica Sheared DNA linear GSS 24-SEP-2001
DEFINITION      ENTTS28TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
                genomic, genomic survey sequence.
ACCESSION      BH164736
VERSION        BH164736.1 GI:15738174
KEYWORDS       GSS.
SOURCE         Entamoeba histolytica
ORGANISM       Entamoeba histolytica
                Eukaryota; Entamoebidae; Entamoeba.
REFERENCE      1 (bases 1 to 856)
AUTHORS        Loftus,B., Wang,Z., Van Aken,S. and Fraser,C.
TITLE          Determination of clone end sequences from Entamoeba histolytica
                HMI:IMSS sheared DNA library (2001)
JOURNAL        Unpublished
COMMENT        Contact: Brendan J Loftus
                Department of Eukaryotic Genomics
                The Institute for Genomic Research
                9712 Medical Center Dr., Rockville, MD 20850, USA
                Tel: 301 838 0208
                Fax: 301 838 3543
                Email: bjloftus@tigr.org
                Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
                DNA library
                Seq primer: M13-Forward
                Class: shotgun
                High quality sequence start: 17
                High quality sequence stop: 629.
FEATURES       Location/Qualifiers
                1..856
                /organism="Entamoeba histolytica"
                /mol_type="genomic DNA"
                /strain="HMI:IMSS"
                /db_xref="taxon:5759"
                /clone_lib="Entamoeba histolytica Sheared DNA"
                /note="Vector: pHOS1; Site 1: Bst I; Constructed at The
                Institute for Genomic Research (TIGR), Rockville, MD.
```

```

Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1999).".

BASE COUNT      347 a      229 c      89 g      191 t
ORIGIN

Alignment Scores:
Pred. No.:      0.0811      Length:      856
Score:          134.00      Matches:      88
Percent Similarity: 33.42%      Conservative: 44
Best Local Similarity: 22.28%      Mismatches: 140
Query Match:      4.12%      Indels:      123
DB:              28      Gaps:      19

US-09-807-802A-15 (1-599) x BH164736 (1-856)

QY  160 TrpGlnArgLeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeu 179
Db  |||  ::  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
9 TGGAGACACCTCATTTGGTAAAT-----TCTTCATTGCTGGA 47

QY  180 PheAsnIleGlnValLysGluValThrThrAsnAspGlyValThr----- 195
Db  ||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
48 TTTAATACTAACCAAAACGTCGGGAACAACATAATCCGTTCAATTCTACTGGAAGTTCAGTC 107

QY  196 ---IleAlaAsnAsnLeuThrSerThrValGlnValPheSerAspSerGluTyrGlnLeu 214
Db  ||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
108 TCAGGAGCAATAAACCCCTTTTGCTACTACTCAA-----AACATAACAACAACT 155

QY  215 ProTyrValLeuGlySerAlaHisGlnGlyCysLeuProPheProAlaAspValPhe 234
Db  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::
156 CCTTTCACAACAGGAACAACATAAC-----CCATTTAAT----- 188

QY  235 MetIleProGlnTyrGlyTyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSer 254
Db  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
189 -----ACTACTAACAAATACTACTACTCA----- 212

QY  255 SerPheTyrCysLeuGluTyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThr 274
Db  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
213 -----ACAACTAATCCATTAAAT 230

QY  275 PheSerTyrThrPheGluGluValProPheHisSerSerTyrAlaHisSerGlnSerLeu 294
Db  ::  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
231 ACTAATACTACAAGCACCACCAATCCATTTAATAACAATACTACTAAT----- 278

QY  295 AspArgLeuMetAsnProLeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsn 314
Db  ||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
279 -----AATCCATTT-----AGTACTAATAAC 299

QY  315 GlnSerGlySerAlaGlnAsnLysAspLeuPheSerArgGlySerProAlaGlyMet 334
Db  ::  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
300 ACTACCAATGGACAACCAAT-----TTAATTAAATCAA---ACTCCTTCTTCTTAAT 347

QY  335 SerValGlnProLysAsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSerLys 354
Db  ::::  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
348 ACTATAACACCCAGGAAAT-----AATACAACTACT 377

QY  355 ThrLysThrAspAsnAsnAsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeu 374
Db  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
378 ACAACAGGAAGTAATCCATTGGTAATTTCACCTACCACCTAAT----- 419

QY  375 AsnGlyArgGluSerIleIleAsnProGlyThrAlaMetAlaSerHisLysAspAspGlu 394
Db  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
420 AATACTACCTCCAGTCAACTTCAACAGGAACACCATCATCACAGGAAGTAATCCATT 479

QY  395 AspLysPhePheProMetSerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSer 414
```





/protein\_id="BAC31343.1"  
/db\_xref="GI:26335285"  
/translation="MADPGMMSLFGE DGS L FSEGLEGE CVYPENPVNPMGQQMPID  
QGFPSLQPSLHHPSPNQNTKLTFEDHYSQYEKMHLMDDQPNRMGSAFGNGLASPHS  
QYHTPPVQVPHGGGGGGMGVYPGIONERHQSFVDGSGMWGRAVQVDPDQIRAPYQ  
QQQPQAPSGPPAQGHPPQHMQMGSYLARGDFSMQHQGPQQQRMGQFSSQGEGLSQGS  
PFIATSGPGLSHMPQQSPSMAPSLRHPVQQQFHHHPAALHGSAHSPRFSNPPTQ  
GAVRPQTLNFSRNQTVPSPTVNSGQYSRYPYSLNQLVNSTGMNQNLGLTNDSTPM  
NQSVPRYPNAVGPSPNSGQGLVHQPIHSSGSLNQMTQTWHPSPQGGTYASPPMSP  
MKAMSNPAGTTPPQVRPGSAGMPMEVGSYPNMPHPQPSHQPPGAMGIGQRMGPRNMQ  
QPRSFMGSSAPRELTHMRPNCGPGVGLADPPQAIQERLI PGQHPGQPSFOQLPTC  
PPLQPHFGLHQSSPPPHHPQWPAQLHPSQNTPKQVPVHQHSPSEPFLEKPVPMQTV  
SAQNAQLVKSDDYLPSEIQQPQKKKKKNHIAAGDSSKGFQKDDFPFGGVENQELRR  
NSLDVSQ"

|   |      |   |  |  |        |               |       |       |  |
|---|------|---|--|--|--------|---------------|-------|-------|--|
| BASE COUNT                                    |      |   |  |  | 542 a  | 686 c         | 541 g | 377 t |  |
| ORIGIN  |      |   |  |  |        |               |       |       |  |
| Alignment Scores:                             |      |   |  |  |        |               |       |       |  |
| Pred. No.:                                    |      |   |  |  | 1.75   | Length:       | 2146  |       |  |
| Score:  |      |   |  |  | 125.50 | Matches:      | 142   |       |  |
| Percent Similarity:                           |      |   |  |  | 32.31% | Conservative: | 69    |       |  |
| Best Local Similarity:                        |      |   |  |  | 21.75% | Mismatches:   | 209   |       |  |
| Query Match:                                  |      |   |  |  | 3.86%  | Indels:       | 237   |       |  |
| DB:   |      |   |  |  | 11     | Gaps:         | 36    |       |  |
| US-09-807-802A-15 (1-599) x AK042727 (1-2146) |      |   |  |  |        |               |       |       |  |
| Qy  | 1    | ThrAlaProGlyLysLysArgProValGluGlnSer                          | -----Pro 13                            |  |        |               |       |       |  |
| Db  | 523  | AGCGCCCCGGGAACGGATTGGCGTCTCCGCACTCACAGTATCACACCCCTCCCGTCCCC   |  |  |        |               |       |       |  |
| Qy  | 14   | GlnGluProAspSerSerSerGlyIleGlyLysThrGly                       | -----GlnGlnProAla 30                   |  |        |               |       |       |  |
| Db  | 583  | CAGGTGCCCCACGGAGGCGGTGTGGCGGTTCAGAT-GGGAGTCTATCCCGGCATCCAGAA  |  |  |        |               |       |       |  |
| Qy  | 31   | LysLys  | -----ArgLeuAsnPheGlyGlnThrGlyAspSer 42 |  |        |               |       |       |  |
| Db  | 642  | CGAAAGGCACGGCAATCTTTTGTGGATGGCGGCTCCATGTGGGCCCCCGCGCGTTCA     |  |  |        |               |       |       |  |
| Qy  | 43   | GluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGly  | 62                                     |  |        |               |       |       |  |
| Db  | 702  | GGTACCAGACCAGATACGAGCCCCCTA-----CCAGCAGCAGCAGCCCGCAGCGCGC     | 752                                    |  |        |               |       |       |  |
| Qy  | 63   | ProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAsp     | 82                                     |  |        |               |       |       |  |
| Db  | 753  | ---TCCTTCGGGGCCCCCGGCACAGGCCACCCCCAGCA-----CATGCAGCAGAT       | 800                                    |  |        |               |       |       |  |
| Qy  | 83   | GlyValGlyAsnAlaSerGlyAsnTrpHis                                | -----CysAspSerThrTrpLeuGly 99          |  |        |               |       |       |  |
| Db  | 801  | -----GGGCAGCTATCTGGCAGCTGGGG-ATTTCTCCATGCAGCAGCAGC            | 844                                    |  |        |               |       |       |  |
| Qy  | 100  | AspArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeu  | 119                                    |  |        |               |       |       |  |
| Db  | 845  | -----GCCAGCCACAGCAGAGGATGGGC-----                             | 867                                    |  |        |               |       |       |  |
| Qy  | 120  | TyrLysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyr  | 139                                    |  |        |               |       |       |  |
| Db  | 868  | -----CAGTTTTCCTCAAGGCCAAGAGGGCCTCAGTCAGGGAAGCCCTTTCATTGGCCACC | 921                                    |  |        |               |       |       |  |
| Qy  | 140  | SerThrProTrpGlyTyrPheAspPhe                                   | -----148                               |  |        |               |       |       |  |
| Db  | 922  | TCAGGACCC---GGCCACCTGTGCAATGCCCCCAGCAGAGCCCCCAGCATGGCCCCCTTCC | 978                                    |  |        |               |       |       |  |
| Qy  | 149  | -----AsnArgPheHisCysHis                                       | -----154                               |  |        |               |       |       |  |
| Db  | 979  | CTGCGCCACCAGTGCAGCAGCAGTTCACCATCACCCCGCTGCTCTCCACGGAGATCC     | 1038                                   |  |        |               |       |       |  |
| Qy  | 155  | -----PheSerProArgAspTrpGlnArgLeuIleAsnAsnAsnTrp               | 168                                    |  |        |               |       |       |  |
| Db  | 1039 | GTTGCCACAGTCCCAGATTCTCCCCCAACCCCTCCTCAACAA-----GGG            | 1083                                   |  |        |               |       |       |  |
| Qy  | 169  | GlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLysGluValThr  | 188                                    |  |        |               |       |       |  |

|    |      |  |      |
|----|------|--|------|
| Db | 1084 | GCTGTCAGGCCACAAACCCCTTAACCTTAGTTCTCGGAACCCAGACAGTCCCCCTCCCTACT | 1143 |
| QY | 189  | ThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGlnValPheSer   | 208  |
| Db | 1144 | GTAACAAC-----TCA   | 1155 |
| QY | 209  | AspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCysLeu         | 226  |
| Db | 1156 | GGGAGTATTCTCGATATCCTTAC-----AGTAACCTAAATCAGGAGTAGTTAACAGT      | 1209 |
| QY | 227  | -----ProPheProAlaAspValPhe                                     | 234  |
| Db | 1210 | ACAGGGATGAATCAAAATTTAGGCCTTACAAACAGTACTCCAATGAATCAGTCGGTA---   | 1266 |
| QY | 235  | MetIleProGlnTyrGlyTyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSer   | 254  |
| Db | 1267 | -----CCAAGATAC-----CCCAATGCCGTGGA-----                         | 1290 |
| QY | 255  | SerPheTyrCysLeuGluTyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThr   | 274  |
| Db | 1291 | -----TTCCCGTCA-----AACAGTGGTCAGGACTCGTG                        | 1320 |
| QY | 275  | PheSerTyrThrPheGluGluValProPheHisSerSerTyrAlaHisSerGlnSerLeu   | 294  |
| Db | 1321 | -----CACCAGCAGCCTATCCACTCCAGTGGCTCACTGAACCAATGAAC              | 1365 |
| QY | 295  | AspArgLeuMetAsnProLeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsn   | 314  |
| Db | 1366 | ACACAAACTATGCACCT-----TCACAGCCTCAGGGA                          | 1398 |
| QY | 315  | GlnSerGlySerAlaGlnAsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMet   | 334  |
| Db | 1399 | ACGTACGCCTCTCCACCTCCCATGTCACCCATGAAAGCAATGAGTAACCCAGCAGGC---   | 1455 |
| QY | 335  | SerValGlnProLysAsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSerLys   | 354  |
| Db | 1456 | ---ACGCCGCTCCGAGGTCCAGGCCCGGAG-TGCTGG-----                     | 1490 |
| QY | 355  | ThrLysThrAspAsnAsnAsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeu   | 374  |
| Db | 1491 | -----GATGCCCATGGAAGTTGGCAG---TTATCCAAATAT-----                 | 1523 |
| QY | 375  | AsnGlyArgGluSerIleIleAsnProGlyThrAlaMetAlaSerHisLysAspAspGlu   | 394  |
| Db | 1524 | ---GCCCCACCTCAGCCATCTCACACGCCCTGGTGGCCATGGGAATCGGACAGAGGAA     | 1580 |
| QY | 395  | AspLysPhePheProMetSerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSer   | 414  |
| Db | 1581 | -----TATGGGCCCCAGAAACATGCAGCAGCC                               | 1607 |
| QY | 415  | AsnThrAlaLeuAspAsnValMetIleThrAsp-----GluGluGluIleLys          | 430  |
| Db | 1608 | TCGTTTCGTTTATGGGCATGTCTCGGCACCCCGAGGAGTGCACCGACACATGAGACCAAA   | 1667 |
| QY | 431  | -----AlaThrAsnProValAlaThrGluArgPhe                            | 440  |
| Db | 1668 | TGGTTGCCCTGGTTCGGCCTTGGCGATCCCAAGCAATCCAGGAACGACTGATACCTGG     | 1727 |
| QY | 441  | GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis   | 460  |
| Db | 1728 | CCAACAGCACCC-----TGGTCAGCAGCCATCTTCCAGCAGTTGCCAACCTGTCC        | 1778 |
| QY | 461  | AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro   | 480  |
| Db | 1779 | TCCACTACAGCCCCACCC-----AGGCCTGCACCA                            | 1808 |
| QY | 481  | IleTrpAlaLysIle-ProHisThrAsp-----GlyHisPheHisProSerProLeuMe    | 498  |
| Db | 1809 | GTCTTCACCTCCACATCCCATCACCAGCCTTGGGCACAACTCCACCCGTCACCC-----    | 1863 |
| QY | 498  | tGlyGlyPheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValPr      | 518  |
| Db | 1864 | -----CAGAATACCCACAGAAAGTG-----CCTGTGCA                         | 1892 |

|    |      |   |       |      |
|----|------|---|-------|------|
| Qy | 518  | oAlaAsnProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTy | Se    | 538  |
|    |      | :::    :::  |       |      |
| Db | 1893 | TCAGCATTTCCATCGGAGCCCTTTCTAGAGAAGCCAGTCCGGATATGACCCAGGT   | CAG   | 1952 |
|    |      | :::    :::  |       |      |
| Qy | 538  | rThrGlyGlnValSerVal-----GluIleGluTrpGluLe                 |       | 550  |
|    |      | :::                         :::                           |       |      |
| Db | 1953 | CGCACAGAACGCTCAGCTAGTGAAGAGCGCAGATTACCTGCCGTCGATAGACGACG  | CAGCC | 2012 |
| Qy | 550  | uGlnLysGluAsnSerLysArgTrpAsnProGluVal                     |       | 562  |
|    |      | :::    :::    :::    :::    :::    :::                    |       |      |
| Db | 2013 | ACAGCAGAAGAAGAAGAAAAAGAAAAACAACCCACATT                    |       | 2049 |

|            |  |         |        |                 |
|------------|--|---------|--------|-----------------|
| RESULT 4   | AK035953   | 2295 bp | linear | HTC 05-DEC-2002 |
| LOCUS      | Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length                        |         |        |                 |
| DEFINITION | enriched library, clone:9630020J18 product:hypothetical protein, full insert sequence. |         |        |                 |

AK035953  
AK035953.1 GI:26084922  
HTC; CAP trapper.  
Mus musculus (house mouse)  
Mus musculus

| SOURCE       | ORGANISM   |
|--------------|--|
| MUS MUSCULUS | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. |

|   |   |
|---|---|
| 1 | Carninci, P. and Hayashizaki, Y.<br>High-efficiency full-length cDNA cloning<br>Meth. Enzymol. 303, 19-44 (1999)<br>99279253<br>10349636  |
| 2 | Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,<br>Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.<br>Normalization and subtraction of cap-trapper-selected cDNAs to<br>prepare full-length cDNA libraries for rapid discovery of new genes<br>Genome Res. 10 (10), 1617-1630 (2000)<br>20499374<br>11042159 |

|  |
|--|
| Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Haxada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. |
| RIKEN integrated sequence analysis (RISA) system-384-format<br>sequencing pipeline with 384 multicapillary sequencer   |
| Genome Res. 10 (11), 1757-1771 (2000)  |
| 20530913   |
| 11076861   |

```

misc_feature      1..2295
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days neonate"
/note="hypothetical protein (evidence:
Truncated-Longest-ORF)"
BASE COUNT      571 a      743 c      586 g      395 t
ORIGIN

Alignment Scores:
Pred. No.:      1.92      Length:      2295
Score:          125.50     Matches:     142
Percent Similarity: 32.31%  Conservative: 69
Best Local Similarity: 21.75%  Mismatches: 209
Query Match:    3.86%      Indels:     237
DR.            11         Gaps:       36

```

|   |        |  |                      |       |
|---|--------|--|----------------------|-------|
| BASE COUNT                                    | 571 a  | 743 c  | 586 g                | 395 t |
| ORIGIN  |        |  |                      |       |
| Alignment Scores:                             |        |  |                      |       |
| Pred. No.:                                    | 1.92   | Length:  | 2295                 |       |
| Score:  | 125.50 | Matches:   | 142                  |       |
| Percent Similarity:                           | 32.31% | Conservative:  | 69                   |       |
| Best Local Similarity:                        | 21.75% | Mismatches:  | 209                  |       |
| Query Match:                                  | 3.86%  | Indels:  | 237                  |       |
| DB:   | 11     | Gaps:  | 36                   |       |
| US-09-807-802A-15 (1-599) x AK035953 (1-2295) |        |  |                      |       |
| Qy  | 1      | ThrAlaProGlyLysLysArgProValGluGlnSer                         | -----Pro 13          |       |
|   |        | :::  |                      |       |
| Db  | 672    | AGGCCCCCGGGAACGGATTGGCGGTCTCCGCACTCACAGTATCACACCCCTCCCGTCCCC | 731                  |       |
| Qy  | 14     | GlnGluProAspSerSerSerGlyIleGlyLysThrGly                      | -----GlnGlnProAla 30 |       |
|   |        |  |                      |       |
| Db  | 732    | CAGGTGCCACCGAGGCGGTGGTGGCGGTCCAGAT-GGGAGTCTATCCCGGCATCCAGAA  | 790                  |       |
|   |        |  |                      |       |



QY 31 LysLys-----ArgLeuAsnPheGlyGlnThrGlyAspSer 42  
Db 791 CGAAGGCACGGGCAATCTTTTGTGGATGGCGGCTCCATGTGGGGCCCCCGGGCGGTTC A 850  
QY 43 GluSerValProAspProGlnProLeuGlyGluProProAlaThrProAlaAlaValGly 62  
Db 851 GGTACCAGACAGATACGAGCCCCCTA-----CCAGCAGCAGCAGCCGCGAGCCGCGC 901  
QY 63 ProThrThrMetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAsp 82  
Db 902 ---TCCTCGGGGCCCCCGGCACAGGGCCACCCCGCA-----CATGCAGCAGAT 949  
QY 83 GlyValGlyAsnAlaSerGlyAsnTrpHis-----CysAspSerThrTrpLeuGly 99  
Db 950 -----GGCAGCTATCTGGCACGTGGG-ATTTCTCCATGCAGCAGCAG- 993  
QY 100 AspArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeu 119  
Db 994 -----GCCAGCCACAGCAGAGGATGGC----- 1016  
QY 120 TyrLysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyr 139  
Db 1017 -----CAGTTTCCCAAGGCCAAGGGCCTCAGTCAGGGAAGCCCTTTCATTGCCACC 1070  
QY 140 SerThrProTrpGlyTyrPheAspPhe----- 148  
Db 1071 TCAGGACCC--GGCCACCTGTGCGACATGCCCCCAGCAGAGCCCCCAGCATGGCCCCCTCC 1127  
QY 149 -----AsnArgPheHisCysHis----- 154  
Db 1128 CTGGGCCACCCAGTGCAGCAGCAGTTCACCATCACCCCGCTGCTCTCCACGGAGATCC 1187  
QY 155 -----PheSerProArgAspTrpGlnArgLeuIleAsnAsnAsnTrp 168  
Db 1188 GTTGCCACAGTCCCAGATTCTCCCCAACCCCTCTCTCAACAA-----GGG 1232  
QY 169 GlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLysGluValThr 188  
Db 1233 GCTGTAGGCCACAAACCCCTTAACCTTAGTTCTCGGAACCAGACAGTCCCCCTCCCTACT 1292  
QY 189 ThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGlnValPheSer 208  
Db 1293 GTAAACAAC-----TCA 1304  
QY 209 AspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCysLeu----- 226  
Db 1305 GGGCAGTATTCTCGATATCCTTAC-----AGTAACCTAATCAGGGATTAGTTAACAGT 1358  
QY 227 -----ProProPheProAlaAspValPhe 234  
Db 1359 ACAGGGATGAATCAAATTTAGGCCTTACAACAGTACTCCATGAATCAGTCGGTA--- 1415  
QY 235 MetIleProGlnTyrGlyTyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSer 254  
Db 1416 -----CCAAGATAC-----CCCAATGCCGTGGGA----- 1439  
QY 255 SerPheTyrCysLeuGluTyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThr 274  
Db 1440 -----TTCCCGTCA-----AACAGTGGTCAGGGACTCGTG 1469  
QY 275 PheSerTyrThrPheGluGluValProPheHisSerSerTyrAlaHisSerGlnSerLeu 294  
Db 1470 -----CACCAGCAGCCTATCCACTCCAGTGGCTCACTGAACCAATGAAC 1514  
QY 295 AspArgLeuMetAsnProLeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsn 314  
Db 1515 ACACAAACTATGCACCCCT-----TCACAGCCTCAGGGA 1547  
QY 315 GlnSerGlySerAlaGlnAsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMet 334  
Db 1548 ACGTAGCCCTCTCCACCTCCCATGTACCCCATGAAGCAATGAGTAACCCAGCAGGC--- 1604

QY 335 SerValGlnProLysAsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSerLys 354  
Db 1605 ---ACGCCGCTCCGAGGTGAGGCCCGGAAG-TGCTGG----- 1639  
QY 355 ThrLysThrAspAsnAsnAsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeu 374  
Db 1640 -----GATGCCCATGGAAGTTGGCAG---TTATCCAAATAT----- 1672  
QY 375 AsnGlyArgGluSerIleAsnProGlyThrAlaMetAlaSerHisLysAspAspGlu 394  
Db 1673 ---GCCCCACCTCAGCCATCTCACCCAGCCCCCTGGTGCCATGGGAATCGGACAGAGAA 1729  
QY 395 AspLysPhePheProMetSerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSer 414  
Db 1730 -----TATGGGCCCCAGAAACATGCAGCAGCC 1756  
QY 415 AsnThrAlaLeuAspAsnValMetIleThrAsp-----GluGluGluIleLys 430  
Db 1757 TCGTTCGTTTATGGGCATGTCTCGGCACCCCGGAGCTGACCCGACACATGAGACCAAA 1816  
QY 431 -----AlaThrAsnProValAlaThrGluArgPhe 440  
Db 1817 TGGTTGCCCTGGTTCGGCCTTGGGATCCACAAGCAATCCAGGAACGACTACCTGG 1876  
QY 441 GlyThrValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHis 460  
Db 1877 CCAACAGCACCC-----TGGTCAGCAGCCATCTTTCCAGCAGTTGCCAACCTGTCC 1927  
QY 461 AlaMetGlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyPro 480  
Db 1928 TCCACTACAGCCCCCACCC-----AGGCCTGCACCA 1957  
QY 481 IleTrpAlaLysIle-ProHisThrAsp-----GlyHisPheHisProSerProLeuMe 498  
Db 1958 GTCTTCACCTCCACATCCCCATCACAGCCTTGGGCACAACTCCACCCGTCACCC----- 2012  
QY 498 tGlyGlyPheGlyLeuLysAsnProProProGlnIleLeuIleLysAsnThrProValPr 518  
Db 2013 -----CAGAATACCCCCACAGAAAGTG-----CCTGTGCA 2041  
QY 518 oAlaAsnProProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSe 538  
Db 2042 TCAGCATTCTCCATCGAGCCCTTTCTAGAGAACCCAGTCCCGGATATGACCCAGGTGAG 2101  
QY 538 rThrGlyGlnValSerVal-----GluIleGluTrpGluLe 550  
Db 2102 CGCACAGAACGCTCAGCTAGTAGAAGAGCGACGATTACCTGCGTCGATAGAGCAGCAGCC 2161  
QY 550 uGlnLysGluAsnSerLysArgTrpAsnProGluVal 562  
Db 2162 ACAGCAGAAGAGAAGAAAAAGAAAAACAACACACATT 2198  
RESULT 5  
BG207226 754 bp mRNA linear EST 21-APR-2001  
LOCUS RST26694 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.  
DEFINITION BG207226  
ACCESSION BG207226  
VERSION BG207226.1 GI:13728913  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 754)  
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,  
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J.,  
Lerner,L., Costanzo,D., McElligott,K., Booser,S., Mays,R., Smith  
E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K., Offenbacher  
J., Danzig,J. and Ducar,M.  
TITLE Creation of genome-wide protein expression libraries using random  
activation of gene expression  
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)  
MEDLINE 21227151

PUBMED 11329013  
COMMENT Contact: Scott J. Cain  
Athersys, Inc.  
3201 Carnegie Ave, Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: scai@atersys.com  
High quality sequence stop: 326.  
Location/Qualifiers  
FEATURES  
source  
1..754  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_line="HT1080"  
/clone\_lib="Athersys RAGE Library"  
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."  
BASE COUNT 223 a 224 c 191 g 115 t 1 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.45 Length: 754  
Score: 125.00 Matches: 50  
Percent Similarity: 33.06% Conservative: 31  
Best Local Similarity: 20.41% Mismatches: 68  
Query Match: 3.84% Indels: 96  
DB: 10 Gaps: 11  
US-09-807-802A-15 (1-599) x BG207226 (1-754)  
QY 307 TyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeuPhe 326  
Db 206 TTTTATCTCAACCAACGGA-----GCAGGAGGTAAATCTAAGGAGCACTGTC 256  
QY 327 SerArgGlySerProAlaGlyMetSerValGlnProLysAsnTlpLeuProGlyProCys 346  
Db 257 CAAAGAAGTACAGAAACAGGCCTGGCCGTGGAATGAGGAACCTGGATG----- 304  
QY 347 TyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThrTrp 366  
Db 305 -----ACTCGACAGCAAGCCGAGAGTCTACAGATGGTAGCATGAACAGCTACAGCTCA 358  
QY 367 ThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThrala 386  
Db 359 CAAGGA-----AATCTGATTTTCCTGGGGTTTCAC 388  
QY 387 MetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIlePhe 406  
Db 389 TTGGCCTCTGATAGCCAGCCAGCCAGCATTCCTGGATGGCCTTGGCCCTGCTCAGCTAGTG 448  
QY 407 GlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAspGlu 426  
Db 449 GGACGCCAGACTCTG----- 463  
QY 427 GluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaValAsn 446  
Db 463 ----- 463  
QY 447 PheGlnSerSerSerThrAspProAlaThrGlyAspValHis---AlaMetGlyAlaLeu 465  
Db 464 -----ACAACACCTGCAATGGGGACATTCACGTTAGGATGATGGCAAC 508  
QY 466 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 485  
Db 509 AAGGGACAGCTG-----GACGCATACTCATCCGGGCC-----GCGCCACTG 550  
QY 486 Pro-----HisThrAspGlyHisPheHis----- 493  
Db 551 CCGAAAAACCAAGCTACAAGACACTGCCACCACCGCTGTCAAAGTAGGCCTACTACAA 610

QY 494 -----ProSerProLeuMetGlyGlyPhe 501  
Db 611 CGGAAGCCGGCCGAGGCAAAAAAACACCCGATGCCGCACCCCTAGCGCCA----- 664  
QY 502 GlyLeuLysAsnProProGlnIleLeuIleLysAsnThr-----ProVal 517  
Db 665 TAAATATCCGACCCACCCACGATCAAGACAAAACCGCACCGGACCGAACCCACG 724  
QY 518 ProAlaAsnProPro 522  
Db 725 CCTGGCAACCTCCG 739  
RESULT 6  
BH152154  
LOCUS  
DEFINITION BH152154 925 bp DNA linear GSS 27-AUG-2001  
ENTPV12TR Entamoeba histolytica Sheared DNA Entamoeba histolytica  
genomic, genomic survey sequence.  
ACCESSION BH152154  
VERSION BH152154.1 GI:15315702  
KEYWORDS GSS.  
SOURCE Entamoeba histolytica  
ORGANISM Entamoeba histolytica  
Eukaryota; Entamoebidae; Entamoeba.  
REFERENCE 1 (bases 1 to 925)  
AUTHORS Loftus,B., Wang,Z., Van Aken,S. and Fraser,C.  
TITLE Determination of clone end sequences from Entamoeba histolytica  
JOURNAL HM1:IMSS sheared DNA library (2001)  
COMMENT Unpublished  
Contact: Brendan J Loftus  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 3543  
Email: bjloftus@tigr.org  
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared  
DNA library  
Seq primer: M13-Reverse  
Class: shotgun  
High quality sequence start: 12  
High quality sequence stop: 720.  
FEATURES  
source  
Location/Qualifiers  
1..925  
/organism="Entamoeba histolytica"  
/mol\_type="genomic DNA"  
/strain="HM1:IMSS"  
/db\_xref="taxon:5759"  
/clone\_lib="Entamoeba histolytica Sheared DNA"  
/note="Vector: PHOS1; Site 1: Bst I; Constructed at The  
Institute for Genomic Research (TIGR), Rockville, MD.  
Genomic DNA isolated from broth cultures of E. histolytica  
using a method described by Clark and Diamond (Clark,  
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a  
method for isolate identification. Exp. Parasitol.  
77:450.). The DNA was mechanically sheared to give a  
tight size distribution (~2 kb). The v + i method used for  
the library construction is described in detail in Smith,  
H.O. and Venter, J.C. (Making small insert libraries for  
whole genome shotgun sequencing projects. In Genome  
Sequencing: A Practical Approach, eds. M. Vaudin and B.  
Barell, Oxford University Press, 1999)."

BASE COUNT 361 a 239 c 96 g 229 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.25 Length: 925  
Score: 121.50 Matches: 86  
Percent Similarity: 32.55% Conservative: 53  
Best Local Similarity: 20.14% Mismatches: 134  
Query Match: 3.74% Indels: 154  
DB: 21 Gaps: 21

US-09-807-802A-15 (1-599) x BH152154 (1-925)

```
QY 112 LeuProThrTyrAsnAsnHisLeuTyrLysGlnLeuSerSerAlaSerThrGlyAlaSer 131
Db 29 CTATCATCTTCGATGAACAT-----CAATTCTCTTCATTTCAATCTGGTCAACAA 79
QY 132 Asn-----AspAsnHisTyrPheGlyTyrSerThrProTyr 143
Db 80 AATCCTTTTGGAACTCCAAACAATACTACACACCTCAATTTGGTACAAATCTTCAATC 139
QY 144 GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTyrGlnArgLeu 163
Db 140 GCTGGATTTAATACTAACCAACGTCGGGAACAACATACTCCG-----181
QY 164 IleAsnAsnAsnTyrGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln 183
Db 182 -----TTCAATTCTACT 193
QY 184 VallysGluValThr---ThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSer 202
Db 194 GGAAGTTCAGTCTCAGGAGCAATAACCTTTTGTCTACTACTCAAAACAATAACAACACT 253
QY 203 ThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHis 222
Db 254 -----CCTTTCACAACAGGAACAACATAAC 277
QY 223 GlnGlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeu 242
Db 278 -----CCATTTAAT-----286
QY 243 ThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhe 262
Db 287 ACTACTAACAATACTACTACTA-----310
QY 263 ProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluVal 282
Db 311 -----ACAACTAATCCATTTAATACTAATACTACAAGCACCACAACAT 352
QY 283 ProPheHisSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIle 302
Db 353 CCATTAAATACAATACTACTAAT-----AATCCATT---385
QY 303 AspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLys 322
Db 386 -----AGTACTAATAACACTACTACCAATGGAACAACCAAT---418
QY 323 AspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeu 342
Db 419 -----TTATTTAATCAA---ACTCCTTCTTCTAATACTATAACACCAGGAAT-----463
QY 343 ProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSer 362
Db 464 -----AATACAACACTACTACAACAGGAAGTAATCCATTGGT 499
QY 363 AsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsn 382
Db 500 AATTTCACTACCCTAAT-----AATACTACCTCCAGTACAACACTCA 541
QY 383 ProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGly 402
Db 542 ACAGGAACAACCATCACACAGGAAGTAATCCATTGGTAAATTCACTCCCTCAAACCTCA 601
QY 403 ValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMet 422
Db 602 GCTCCAACAACAGGAATAATAACACAGGAACAACACTACCACAACAGGAATAATAACACT 661
QY 423 IleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThr 442
Db 662 TCAACA-----GGAAGTAATCCA-----TTTGGT---685
QY 443 ValAlaValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMet 462
Db 686 -----AATTTCACTACTTCAAACTCAACTACCACAACAGGAACGACAGCTACACA 736
```

```
QY 463 GlyAlaLeuProGlyMetValTrpGlnAspArgAsp-ValTyrLeuGlnGlyProIleTr 482
Db 737 GGA-----AATGACACAACACTACAGGGACCA-----761
QY 482 pAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheG1 502
Db 762 -----CCCCACACACAGGGGAGTATCCAT-----TTGGGTAATTTC--797
QY 502 yLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProPr 522
Db 798 -ATACAATAATAACCACTTTCTATGTAGTACAACTCCCAACACACAGGAAGTAACCCACT 856
QY 522 oAlaGluPheSerAlaThr 528
Db 857 TGGGTATTTCACCTCA 875

RESULT 7
BH115587 753 bp DNA linear GSS 19-JUL-2001
LOCUS RPCI-24-358F16.TV RPCI-24 Mus musculus genomic clone RPCI-24-358F16
DEFINITION , genomic survey sequence.
ACCESSION BH115587
VERSION BH115587.1 GI:14954954
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 753)
AUTHORS Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
Russell,D., de Jong,P. and Fraser,C.M.
TITLE Mouse BAC End Sequences from Library RPCI-24
JOURNAL Unpublished
COMMENT Other_GSSs: RPCI-24-358F16.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC
library availability, please contact Pieter de Jong
(pdejong@mail.cho.org). Clones may be purchased from BACPAC
Resources (http://www.choi.org/bacpac/orderingframe.htm). BAC end
plate: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 358 row: F column: 16
Seq primer: T7
Class: BAC ends.
FEATURES
source Location/Qualifiers
1..753
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-358F16"
/sex="Male"
/cell_type="Spleen/Brain"
/clone_lib="RPCI-24"
/note="Vector: pTARBAC1; Site 1: BamH1; Site 2: BamH1;
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pTARBAC1 cloning vector at the
BamH1 sites using MboI partially digested male C57BL/6J
DNA."
BASE COUNT 211 a 178 c 167 g 197 t
ORIGIN
Alignment Scores: 3.31 Length: 753
Pred. No.: 115.50 Matches: 44
Score: 42.22% Percent Similarity: 13
```



Best Local Similarity: 32.59% Mismatches: 48  
Query Match: 3.55% Indels: 32  
DB: 28 Gaps: 4

US-09-807-802A-15 (1-599) x BH115587 (1-753)

```
QY 2 AlaProGlyLysLys-----ArgProValGluGlnSerPro----- 13
Db 348 GCTCCTGGGAAGAAATAGATTCCGACCTTA-AACTCCTCCCCCATGAAGAAGTACTACT 406
QY 14 -----GlnGluProAspSerSerSerGlyIleGlyLysThrGlyGlnProAla 30
Db 407 CGCTCTTTTCAGAGTTATGCCAAAACAGTAGAAGTGGAGGCTGGCTCATTAGTCTAT 466
QY 31 LysLysArgLeuAsnPheGlyGlnThrGlyAspSerGluSerValProAspProGlnPro 50
Db 467 GACAACAGATGTAGCAGTTTCCAGCAGATATCTCCAATTTA----- 508
QY 51 LeuGlyGluProProAlaThrProAlaAlaValGlyProThrThrMetAlaSerGly--- 69
Db 509 -----GACACTTCTATAATCTCTGGAGCT 532
QY 70 GlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGlyAsnAlaSerGly 89
Db 533 GGAGGTCTACTAATGGGCAACAACACAGAC-ACTGATGGAGTGAGCAATGCCATGGAT 591
QY 90 AsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThrSerThrArgThr 109
Db 592 GATTGGCATTGCCATTCCAAGTGTATGGTAGACTGAGTCAATTACCTAATCTCCCCAGACC 651
QY 110 TrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSer 124
Db 652 TGGGTCTCTGCCCACTCCAACCTAACAATTTCCACAACACTATATGAAC 696
```

RESULT 8  
BQ642977  
LOCUS  
DEFINITION BQ642977 1042 bp mRNA linear EST 15-JUL-2002  
AGENCOURT\_8289156 NIH\_MGC\_99 Homo sapiens cDNA clone IMAGE:6303333  
5', mRNA sequence.

ACCESSION BQ642977  
VERSION BQ642977.1 GI:21767149  
KEYWORDS EST.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1042)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished

REFERENCE  
AUTHORS Contact: Robert Strausberg, Ph.D.  
TITILE Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
JOURNAL Tissue Procurement: Lou Staudt  
COMMENT cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LLCM2522 row: d column: 22  
High quality sequence stop: 481.  
Location/Qualifiers

FEATURES  
source  
1..1042  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6303333"  
/tissue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_99"  
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:

GGCACGAG(G). Size-selected >500bp for average insert size  
1.8kb. Library constructed by Ling Hong in the laboratory  
of Gerald M. Rubin (University of California, Berkeley)  
using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
II RT (Life Technologies). Note: this is a NIH\_MGC  
Library."

BASE COUNT 118 a 412 c 308 g 204 t  
ORIGIN

Alignment Scores:

Pred. No.: 6.42 Length: 1042  
Score: 114.50 Matches: 79  
Percent Similarity: 33.23% Conservative: 31  
Best Local Similarity: 23.87% Mismatches: 118  
Query Match: 3.52% Indels: 104  
DB: 13 Gaps: 19

US-09-807-802A-15 (1-599) x BQ642977 (1-1042)

```
QY 2 AlaProGlyLysLysArgProValGluGlnSer-----ProGlnGlu 15
Db 107 TCACCCGGGCTACGGCGGCCGCGCAGGGGGCAGTGGCGGCCGTTGTCTAGGGCCC----- 160
QY 16 ProAspSerSerSerGlyIleGlyLysThrGlyGln---GlnProAlaLysLysArgLeu 34
Db 161 -----GGAGGTGGGGCGCGCGCTCGGGCGCTACGAACCCGCGAGGCCACGCTT 211
QY 35 AsnPheGlyGlnThr-----GlyAsp----- 41
Db 212 GGCTCGCGCCGGGTGCGGGTGAGGCCATGGCGCAGCCGCGCAGCTCCCGCGCGGAGG 271
QY 42 SerGluSerValProAspProGlnProLeuGlyGluProAlaThr----- 57
Db 272 ACAGGTGCGCGCCCTCGAGCCACGGCGCGCGCGCGCTGCGGCGGGAAGAGAGGTG 331
QY 58 ---ProAlaAlaValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAsp 76
Db 332 CGGAGCGCCGCTGCGGGCCCCGACGTGGAGCCCGCGGGGACGCCAGCCCCGCCGCC 391
QY 77 AsnAsnGluGlyAlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThr 96
Db 392 AACAGGACGGAGACGCCGCGCGTGGGC-----AGCGGCCACTGGGAGCTGAGGTGCCAT 445
QY 97 TrpLeuGlyAspArgValIleThrThrSerThrArgThrTrpAlaLeuProThrTyrAsn 116
Db 446 CGCCTGCAGGATTCTTTATTACCTCTGACAGT-----GGCCTTCACAACATACCGT 496
QY 117 AsnHisLeuTyrLysGlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyr 136
Db 497 GGCCAT-----CCTGAACCTGGCTGTCCGCGGACTCTCCCCCTCTCTACACCCG 544
QY 137 PheGlyTyrSer-----ThrProTrpGlyTyrPheAspPheAsnArgPheHis 152
Db 545 CCGGCTTACTCCCCGAAGAAACCTCTCCA-----CAT 577
QY 153 CysHisPheSerProArgAsp----- 159
Db 578 ATTCACTTCCCTCCCGCTCCCGCTCCCATGGTCCCCCCCCCTGCTGCTGCTGCTGCT 637
QY 160 -----TrpGlnArgLeuIleAsnAsn 166
Db 638 CTCGTATCCCGGCTTGCTCCTCACCGGGTCGACCGTTGTGGATGCCCTC----- 688
QY 167 AsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLysGlu 186
Db 689 -----CCCCCGCTTACACTGTTCCGGAATTTCTCTCTCTCCCGCGCGGTATT 733
QY 187 ValThrThrAsnAspGlyVal-----ThrThrIleAlaAsnAsnLeuThr 201
Db 734 CTTGTGTGCGCATGCTCGCCTCCGACTTCGTGCGCCTTCCACTTGCCTCCGCTTCCCTCCTT 793
QY 202 SerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGly---Ser 220
```

|            |     |  |   |                 |
|------------|-----|--|---|-----------------|
| Db         | 794 | ---CGCGCGTGGTGGTGGCTCTGCGACATTTC                                     | CGGTTCCGCTAGCTCGCCCTCC                    | 850             |
| Qy         | 221 | AlaHisGlnGlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGly            |   | 240             |
| Db         | 851 | CCCTCCTTGGGTGGCTCCTCCCTGTC   | CCCTCGGTCGTCGCTCCG-----                   | 901             |
| Qy         | 241 | TyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGlu         |   | 260             |
| Db         | 902 | TACCG-GCTCTCAACACTTCTTCGCC   | -----CGCGTCCCTAATATTGCCTGGAG              | 951             |
| Qy         | 261 | -----TyrPheProSerGlnMetLeu   | 267                                       |                 |
| Db         | 952 | GTTATTGCACGCTTCTTCCCTCCCGTCCCATC                                     | 984                                       |                 |
| RESULT 9   |     |  |   |                 |
| AK076994   |     |  |   |                 |
| LOCUS      |     |  |   |                 |
| DEFINITION |     | AK076994   | 5809 bp mRNA linear                       | HTC 07-DEC-2002 |
|            |     | Mus musculus adult male testis cDNA, RIKEN full-length enriched      |   |                 |
|            |     | library, clone:4931406P16  | product:hypothetical protein, full insert |                 |
|            |     | sequence.  |   |                 |
| ACCESSION  |     | AK076994.1   | GI:26345807                               |                 |
| VERSION    |     | HTC; CAP trapper.  |   |                 |
| KEYWORDS   |     | Mus musculus (house mouse)   |   |                 |
| SOURCE     |     | Mus musculus   |   |                 |
| ORGANISM   |     | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;    |   |                 |
|            |     | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.   |   |                 |
| REFERENCE  |     | 1  |   |                 |
| AUTHORS    |     | Carninci,P. and Hayashizaki,Y.                                       |   |                 |
| TITLE      |     | High-efficiency full-length cDNA cloning                             |   |                 |
| JOURNAL    |     | Meth. Enzymol. 303, 19-44 (1999)                                     |   |                 |
| MEDLINE    |     | 99279253   |   |                 |
| PUBMED     |     | 10349636   |   |                 |
| REFERENCE  |     | 2  |   |                 |
| AUTHORS    |     | Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,        |   |                 |
|            |     | Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.       |   |                 |
|            |     | Normalization and subtraction of cap-trapper-selected cDNAs to       |   |                 |
|            |     | prepare full-length cDNA libraries for rapid discovery of new genes  |   |                 |
| TITLE      |     | Genome Res. 10 (10), 1617-1630 (2000)                                |   |                 |
| JOURNAL    |     | 20499374   |   |                 |
| MEDLINE    |     | 11042159   |   |                 |
| PUBMED     |     | 3  |   |                 |
| REFERENCE  |     | Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,  |   |                 |
|            |     | Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,    |   |                 |
|            |     | Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,    |   |                 |
|            |     | Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,   |   |                 |
|            |     | Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,   |   |                 |
|            |     | Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,  |   |                 |
|            |     | Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.       |   |                 |
|            |     | RIKEN integrated sequence analysis (RISA) system--384-format         |   |                 |
|            |     | sequencing pipeline with 384 multicapillary sequencer                |   |                 |
| TITLE      |     | Genome Res. 10 (11), 1757-1771 (2000)                                |   |                 |
| JOURNAL    |     | 20530913   |   |                 |
| MEDLINE    |     | 11076861   |   |                 |
| PUBMED     |     | 4  |   |                 |
| REFERENCE  |     | Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,   |   |                 |
|            |     | Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,   |   |                 |
|            |     | Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,   |   |                 |
|            |     | Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,   |   |                 |
|            |     | Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,        |   |                 |
|            |     | Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,       |   |                 |
|            |     | Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,                |   |                 |
|            |     | Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,      |   |                 |
|            |     | Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,        |   |                 |
|            |     | Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,          |   |                 |
|            |     | Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,              |   |                 |
|            |     | Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,        |   |                 |
|            |     | Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,                |   |                 |
|            |     | Marchionni,L., Mashima,J., Mazzairelli,J., Mombaerts,P., Nordone,P., |   |                 |
|            |     | Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,          |   |                 |
|            |     | Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,  |   |                 |
|            |     | Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,          |   |                 |
|            |     | Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.    |   |                 |

|           |  |   |  |  |
|-----------|--|---|--|--|
| TITLE     |  | and Hayashizaki,Y.  |  |  |
| JOURNAL   |  | Functional annotation of a full-length mouse cDNA collection        |  |  |
| MEDLINE   |  | Nature 409 (6821), 685-690 (2001)                                   |  |  |
| PUBMED    |  | 21085660  |  |  |
| REFERENCE |  | 11217851  |  |  |
| AUTHORS   |  | 5   |  |  |
|           |  | The FANTOM Consortium and the RIKEN Genome Exploration Research     |  |  |
|           |  | Group Phase I & II Team.  |  |  |
|           |  | Analysis of the mouse transcriptome based on functional annotation  |  |  |
|           |  | of 60,770 full-length cDNAs   |  |  |
| TITLE     |  | Nature 420, 563-573 (2002)  |  |  |
| JOURNAL   |  | 6 (bases 1 to 5809)   |  |  |
| REFERENCE |  | Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,     |  |  |
|           |  | Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,          |  |  |
|           |  | Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,           |  |  |
|           |  | Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,       |  |  |
|           |  | Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,      |  |  |
|           |  | Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,          |  |  |
|           |  | Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,      |  |  |
|           |  | Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,         |  |  |
|           |  | Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,        |  |  |
|           |  | Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,           |  |  |
|           |  | Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A.,         |  |  |
|           |  | Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.             |  |  |
|           |  | Direct Submission   |  |  |
| TITLE     |  | Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of     |  |  |
| JOURNAL   |  | Physical and Chemical Research (RIKEN), Laboratory for Genome       |  |  |
|           |  | Exploration Research Group, RIKEN Genomic Sciences Center (GSC),    |  |  |
|           |  | RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, |  |  |
|           |  | Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,        |  |  |
|           |  | URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,             |  |  |
|           |  | Fax:81-45-503-9216)   |  |  |
|           |  | cDNA library was prepared and sequenced in Mouse Genome             |  |  |
| COMMENT   |  | Encyclopedia Project of Genome Exploration Research Group in Riken  |  |  |
|           |  | Genomic Sciences Center and Genome Science Laboratory in RIKEN.     |  |  |
|           |  | Division of Experimental Animal Research in Riken contributed to    |  |  |
|           |  | prepare mouse tissues.  |  |  |
|           |  | Please visit our web site for further details.                      |  |  |
|           |  | URL:http://genome.gsc.riken.go.jp/                                  |  |  |
|           |  | URL:http://fantom.gsc.riken.go.jp/.                                 |  |  |
| FEATURES  |  | Location/Qualifiers   |  |  |
| source    |  | 1..5809   |  |  |
|           |  | /organism="Mus musculus"  |  |  |
|           |  | /mol_type="mRNA"  |  |  |
|           |  | /strain="C57BL/6J"  |  |  |
|           |  | /db_xref="FANTOM_DB:4931406P16"                                     |  |  |
|           |  | /db_xref="MGI:1910824"  |  |  |
|           |  | /db_xref="taxon:10090"  |  |  |
|           |  | /clone="4931406P16"   |  |  |
|           |  | /sex="male"   |  |  |
|           |  | /tissue_type="testis"   |  |  |
|           |  | /clone_lib="RIKEN full-length enriched mouse cDNA library"          |  |  |
|           |  | /dev_stage="adult"  |  |  |
|           |  | 415..3594   |  |  |
|           |  | /note="unnamed protein product; hypothetical protein                |  |  |
|           |  | (evidence: rCDS,ProCrest,decoder,NCBI CDS                           |  |  |
|           |  | Predictor,Longest-ORF)  |  |  |
|           |  | putative"   |  |  |
|           |  | /codon_start=1  |  |  |
|           |  | /protein_id="BAC36555.1"  |  |  |
|           |  | /db_xref="GI:26345808"  |  |  |
|           |  | /db_xref="MGI:1924311"  |  |  |
|           |  | /translation="MYCCSAQDNKMDYKRRFLLGSKQKVQHQOQYMPPELGRALSA            |  |  |
|           |  | LASTAPSAPLGSLTAAGSCHHAMPHSTPIADIQQGISKYLDAIWNVFCRASTFLTDLFS         |  |  |
|           |  | TVFRNSHYSKAAMQLKDQVEHVMEAAASRLTSAIKPEIAQMLMELSAGANFTDQKEFS          |  |  |
|           |  | LQDIEVLGRCFLTVQVHFQFLTHALQKVQVVAHSCFAEVVPEKQNSLSGMGHTPE             |  |  |
|           |  | LEEAVRSWRGAEEATSRRLRERGCDGCLAGIEVQQLFCQSAAIPEHQLKELNKIDSA           |  |  |
|           |  | LQAYKIALESGLGHCYAMKAGFLNPKAIEASLQGCCSEAEAAQQTGRRQTPOPMQCE           |  |  |
|           |  | LPTVPVQIGSHFLKGVSFNESAADNLKLKTHMLQLIKEAGCYNGITPRDDLPTVEVL           |  |  |
|           |  | NOVCPSMWRGACKTAVQLLFGQAGLVVVDPAQIENKEAYAPQISLEGSRLVVQVPSTW          |  |  |
|           |  | CLKEDPATMSLLQRSLDPEKTLGLVDLVLYTAVLDLNRWRTGREQALPCIQIQLDIC           |  |  |
|           |  | DFGNQADLPSSGSKSSGLQKTFSKLTSRFTKVKVCTNSGSTHYSKNIFTAGCSEERA           |  |  |
|           |  | KMMGNIDTRLQSIILNIGNPRTTDPSSQSSQSSNPVMVNGFLLERENFLHGGDDGKDE          |  |  |









Db 436 AAAGGCAGATTGTTGCCAGTCAGCTAGAAAGATGTAGGCTTGGAGCAGAAATCGCCCAAGCA 495  
QY 166 nAsnTrpGlyPheArgProLysArgLeuAsnPhelYsLeuPheAsnIleGlnValLys-G 186  
Db 496 T-----CGCCAGCACCAGCTCAACTGAAAGTCAATTTCCTTGGGATCAACAG 543  
QY 186 luValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGlnV 206  
Db 544 ATGTG---CCAAATCCTGGTGTG-----AGCAAACCTAGAG 576  
QY 206 alPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer---AlaHisGlnGlyC 225  
Db 577 TT---TCTGACACTATTCAATCCCAACTCAAGAGTTCA-----ACAC 705  
QY 225 ysLeuProPropheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuA 245  
Db 634 CCCTC-----TACTCACAGAAACAGACTTCTCTCCATGAAA 669  
QY 245 snAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerG 265  
Db 670 GTGAGGGATCGTTGGGAAACTCAAGAAGTTCA-----ACAC 705  
QY 265 lnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPropheH 285  
Db 706 AGATG-----AATTCTTACTCTGACAGTGGGTACCAGGAAGCAGGAGTTTCC 753  
QY 285 isSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnT 305  
Db 754 ACAACAGCCAAACCGTGAACAAGGCA--GACAGCAGACAGCATCCATTTCACAGGATCAA 810  
QY 305 yrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuL 325  
Db 811 CAAGTAACCATGTGTGGTAGGACTTCAAGAGCTGAAGACAAACACTGGTTTCAGCCATCAG 870  
QY 325 euPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyP 345  
Db 871 TAGCCAATCGGGCCATCGGAGAGTTAGTTTCAGTTCCATCTAGAGCAGACAGTCTCCTTCT 930  
QY 345 roCysTyrArgGlnGlnArgValSer---LysThrLysThrAspAsnAsnAsnSerAsnp 364  
Db 931 A-TGTTACCAGCAGACAGCGGTGTCTCTTCAAGGGGTCACTGAGAACTTCTCTGGGTAGT 989  
QY 364 heThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProG 384  
Db 990 GGATTGGCTCTCCATCAGT-----GACGGACTCCCGACC---TCTGAACCCC 1034  
QY 384 lyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValM 404  
Db 1035 AGTGCTTACTCTCTCCAGCAC-----1054  
QY 404 etIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleT 424  
Db 1055 -----GCTCCCTGCACAGCGGGCGCCTCTCCATACTCA 1088  
QY 424 hrAspGluGluGluIleLysAlaThrAsnProvalAlaThrGluArgPheGlyThrValA 444  
Db 1089 CA-----GAGACCCGCTCCCAACAGCCGTCGCGGGTCCGGTCTGTCTCACCTCC 1139  
QY 444 laValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyA 464  
Db 1140 CGGCAGACCTCCAATCCCAACGGACCAAGTCCCTCAGTACCAGACCACCAGGGTGGG 1199  
QY 464 la-----464  
Db 1200 TCCCCACTGACCCCTGACTGATGCCCCAGACTCGAGTAGTCTCCCATCCCAAGCCAGGTG 1259  
QY 465 -----LeuProGlyMetValTrpGlnAspArg-----473  
Db 1260 GGGTCATCGTCCCGAAACCGTCTGGGATGACCGCGTACCAGACACCTGGGACCTTCA 1319  
QY 474 -----AspValTyrLeuGlnGlyProIleTrp---AlaLysIleProHis-----T 488

Db 1320 CTGCAAAAGGAGTGTTCATGACATGGACCAATTTGGACAGCAGCATGATGACATTTATGAA 1379  
QY 488 hrAspGlyHisPhe-HisProSerProLeuMetGly-----499  
Db 1380 AGGATGGTTCCACCTCGGCCAGACAGCCTGACAGGCTTACGGAGTTCTTATGCCAGTCAG 1439  
QY 500 -----GlyPheGlyLeuLysAsnPro---ProProGlnIleLeuIleLysAsn 514  
Db 1440 CATAGTCAGCTTGGGCAAGAGCTCCGTTTCAGCTGTGTCTCTGACTTGACATC-----1493  
QY 515 ThrProVal-----ProAlaAsnPro 521  
Db 1494 ACTCCTATCTATGAAGGAGGACCTACTACAGTCCAGTGTACCGCAGCCCAAAACCACGGG 1553  
QY 522 ProAlaGluPheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGln 541  
Db 1554 ACCGTGGAGCTCCAGGGGTACACAGACGCA-----TTGTATCGCACAGGCTCA 1601  
QY 542 ValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGlu 561  
Db 1602 GTAGGTATT-----GGAAATCTACAAAGGACATCCAGCCAGCGA-----1640  
QY 562 ValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsn 581  
Db 1641 -----AGCACCTTACATACCAAGAAAT-----AACTATGCCCTGAACACAGCA 1685  
QY 582 GlyLeuTyrThrGluPro---ArgProIleGlyThrArg 593  
Db 1686 GCTACCTATCGGAGGCCCTACAGGCCAGTTCAGTACCGA 1724  
RESULT 12  
AY103595  
LOCUS AY103595 1960 bp mRNA linear HTC 16-OCT-2002  
DEFINITION Zea mays PCO108378 mRNA sequence.  
ACCESSION AY103595  
VERSION AY103595.1 GI:21206673  
KEYWORDS HTC.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 1960)  
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,  
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.  
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of  
Overgo Probes  
JOURNAL Unpublished (2002)  
REFERENCE 2 (bases 1 to 1960)  
AUTHORS Coe,E.H.  
TITLE Direct Submission  
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of  
Missouri, Columbia, MO 65211, USA  
COMMENT If you are interested in getting corresponding physical clones,  
these are publicly available from ZmDB and may be found by BLAST  
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,  
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the  
maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
Schnable, Iowa State, then clones may be requested from ZmDB:  
www.zmdb.iastate.edu.  
FEATURES  
Location/Qualifiers  
1..1960  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="MaizeDB:636289"  
/db\_xref="taxon:4577"  
/clone\_lib="Maize Mapping Project/DuPont Consensus  
Library"  
/note="this sequence is part of a project of EST  
assemblies resulting from the application of public  
contigs to seed DuPont contigs; this resource was  
assembled by DuPont as part of a collaboration for the  
overgo addressing of BACs in conjunction with the Maize



| BASE COUNT                                    |     |   |     | Mapping Project" |  |  |  | 497 a 461 c 489 g 513 t |  |  |  |
|---|-----|---|-----|------------------|--|--|--|-------------------------|--|--|--|
| ORIGIN  |     |   |     |                  |  |  |  |                         |  |  |  |
| Alignment Scores:                             |     |   |     |                  |  |  |  |                         |  |  |  |
| Pred. No.:                                    |     |   |     | 21.3             |  |  |  | Length: 1960            |  |  |  |
| Score:  |     |   |     | 113.00           |  |  |  | Matches: 117            |  |  |  |
| Percent Similarity:                           |     |   |     | 31.02%           |  |  |  | Conservative: 75        |  |  |  |
| Best Local Similarity:                        |     |   |     | 18.90%           |  |  |  | Mismatches: 235         |  |  |  |
| Query Match:                                  |     |   |     | 3.48%            |  |  |  | Indels: 192             |  |  |  |
| DB:   |     |   |     | 11               |  |  |  | Gaps: 24                |  |  |  |
| US-09-807-802A-15 (1-599) x AY103595 (1-1960) |     |   |     |                  |  |  |  |                         |  |  |  |
| QY  | 17  | AspSerSerGlyIleGlyLysThrGlyGlnGlnProAlaLysLysArgLeuAsnPhe     | 36  |                  |  |  |  |                         |  |  |  |
| Db  | 34  | GACAGTCTGAGGCGCTGCGGGCAGTCTAACAGCACCCCTCTCTCGTTCGCACTC        | 87  |                  |  |  |  |                         |  |  |  |
| QY  | 37  | GlyGlnThrGlyAspSerGluSerValProAspProGlnProLeuGlyGluProAla     | 56  |                  |  |  |  |                         |  |  |  |
| Db  | 88  | ---CGTTCGTCTGACATCTCTCCGTCCTTTCTCTTCTGAGGCTCGCGAACCCCGACAA    | 144 |                  |  |  |  |                         |  |  |  |
| QY  | 57  | ThrProAlaAlaValGlyProThrThrMetAlaSerGlyGlyGlyAlaProMetAlaAsp  | 76  |                  |  |  |  |                         |  |  |  |
| Db  | 145 | TGGCCGCAACCCGCGGTGTCGGTCGACGAGAGCTCGACAAAGCTTCGCGCCGAGGTCGCCA | 204 |                  |  |  |  |                         |  |  |  |
| QY  | 77  | AsnAsnGluGlyAlaAspGlyValGlyAsnAlaSerGlyAsnTrpHisCysAspSerThr  | 96  |                  |  |  |  |                         |  |  |  |
| Db  | 205 | AACTCAGCCAGATCAGCGAGAACGAGAGCGCGGTTTCATCAGCCTCGTGTACGCTACC    | 264 |                  |  |  |  |                         |  |  |  |
| QY  | 97  | TrpLeuGlyAspArgValIleThrThrSerThrArgThr                       | 109 |                  |  |  |  |                         |  |  |  |
| Db  | 265 | TCAGTGGGAGGCGGAGCAGATCGAGTGGAGCAAGATCCAGACCCCGACCGATGAGGTAG   | 324 |                  |  |  |  |                         |  |  |  |
| QY  | 110 | TrpAlaLeuProThrTyrAsnAsnHisLeuTyrLys                          | 121 |                  |  |  |  |                         |  |  |  |
| Db  | 325 | TGGTCCGTACGATACCTCAGCTCGCTCCTCTGAAGATCTCGAGGAGACGAAGAGCTGC    | 384 |                  |  |  |  |                         |  |  |  |
| QY  | 122 | -----GlnIleSerSerAlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGly      | 138 |                  |  |  |  |                         |  |  |  |
| Db  | 385 | TGGACAAGCTCGTTGTGCTCAAGCTCAACGGAG                             | 420 |                  |  |  |  |                         |  |  |  |
| QY  | 139 | TyrSerThrProTrpGly-TyrPheAspPheAsnArgPheHisCysHisPheSerProAr  | 158 |                  |  |  |  |                         |  |  |  |
| Db  | 421 | TCGGGACGACCATGGGCTGCAC  | 471 |                  |  |  |  |                         |  |  |  |
| QY  | 158 | g-----AspTrpGlnArgLeu-IleAsnAsnAsnTrpG                        | 169 |                  |  |  |  |                         |  |  |  |
| Db  | 472 | ATGGGTTACATTCTCTTGACCTTATTGTGATTCAAAATCGAGTCCCTGAAACAAGAGTATG | 531 |                  |  |  |  |                         |  |  |  |
| QY  | 169 | lyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLysGluValThrT  | 189 |                  |  |  |  |                         |  |  |  |
| Db  | 532 | GATGTAATGTCCCTTTACTTCTGATGAACCTTTTCAAT                        | 570 |                  |  |  |  |                         |  |  |  |
| QY  | 189 | hrAsnAspGlyValThrThrIleAla-----AsnAsnLeuThrSerThrValGlnV      | 206 |                  |  |  |  |                         |  |  |  |
| Db  | 571 | CCCATGATGACACACAGAGATCGTTGAGAAAGTATCCAACTCCAAACATTGAAATTCATA  | 630 |                  |  |  |  |                         |  |  |  |
| QY  | 206 | alPheSerAspSerGluTyr-----GlnLeuProTyrV                        | 217 |                  |  |  |  |                         |  |  |  |
| Db  | 631 | CTTTCAATCAGAGCCAGTATCTCTCGCATTTGTTACGAGGACTTCTTGCCACTTCCAAGCA | 690 |                  |  |  |  |                         |  |  |  |
| QY  | 217 | alLeuGlySerAlaHisGlnGlyCysLeuProPheProAlaAspValPheMetIleP     | 237 |                  |  |  |  |                         |  |  |  |
| Db  | 691 | AAGGGAATCTGGGAAGGATGGCTGGTATCTCTCAGGCCATGGTGTGTTCCCC          | 746 |                  |  |  |  |                         |  |  |  |
| QY  | 237 | roGlnTyrGlyTyrLeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheT  | 257 |                  |  |  |  |                         |  |  |  |
| Db  | 747 | -----TCTTTGAATAACAGTGGAAACTC-----                             | 770 |                  |  |  |  |                         |  |  |  |
| QY  | 257 | yrCysLeuGluTyrPheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerT  | 277 |                  |  |  |  |                         |  |  |  |
| Db  | 771 | -----GACATCTTATTGGCTCAGGGCAAGGAGTATGTCTTC-----                | 806 |                  |  |  |  |                         |  |  |  |

|    |      |  |      |
|----|------|--|------|
| Qy | 277  | yrThrPheGluGluValPropheHisSerSerTyrAlaHisSerGlnSerLeuAspArgL | 297  |
| Db | 807  | -----GTTGCTAACTCAGACAACTTGGGTGCTA                            | 834  |
| Qy | 297  | euMetAsnProLeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerG | 317  |
| Db | 835  | TAGTCGACATCAAGATC-----CTGAACCATCTGATCAATAACAGAAATGAA         | 881  |
| Qy | 317  | lySerAlaGlnAsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValG | 337  |
| Db | 881  | -----  | 881  |
| Qy | 337  | lnProLysAsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysT | 357  |
| Db | 882  | -----TACTGCATGGAGGTTACTCCAAAAACATTTGG                        | 912  |
| Qy | 357  | hrAspAsnAsnAsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyA | 377  |
| Db | 913  | CTGATGTTAAA-----GGCGGTACTCTCATCTCTTACGAAGGAA                 | 951  |
| Qy | 377  | rgGluSerIleIleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysP | 397  |
| Db | 952  | GAGTTCAGCTTTTGGAGATTGCCCAAGTACCTGATGAGCATGTGAATGAG-----T     | 1002 |
| Qy | 397  | hePheProMetSerGlyValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrA | 417  |
| Db | 1003 | TAAATCAATCGAGAAAGTTTAAAGATATTCAACACTAACAACTTGTGGTGAACCTTAAAG | 1062 |
| Qy | 417  | laLeuAspAsnVal-----MetIleThrAspGluGluG                       | 428  |
| Db | 1063 | CTGTCAAGAGACTAGTAGAGGCTGAGGCACCTTAAGATGGAAATATTCCAAACCCCAAGG | 1122 |
| Qy | 428  | luIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaValAsnPheG | 448  |
| Db | 1123 | AAGTTGATGGTGTGAAAGTCTTCAACTTGAACCTGCAGCTGGTGCAGCTATTCTGTTTC  | 1181 |
| Qy | 448  | lnSerSerSerThrAspProAlaThrGly---AspValHisAlaMetGlyAlaLeuProG | 467  |
| Db | 1182 | -----TTTGACAAAGCGATTGGAATTAATGTTCCCGCTCAAGATTCTCCCGG         | 1230 |
| Qy | 467  | lyMetValTrpGlnAsp-----ArgAspValTyrLeuGlnGlyProIleT           | 482  |
| Db | 1231 | TGAGGCTACATCTGATTTATTGCTTGTGCAGTCTGATCTTAC-----              | 1274 |
| Qy | 482  | rpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheG | 502  |
| Db | 1275 | -----ACCTTGGTTGATGGCTTTG                                     | 1293 |
| Qy | 502  | lyLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProp    | 522  |
| Db | 1294 | TCATCCGCAATCCA-----TCCAGAGCGAAATCCAGCTAACCCCTT               | 1332 |
| Qy | 522  | roAlaGluPheSerAla-----ThrLysPheAlaSerPheIleThrGlnTyrSerThrG  | 540  |
| Db | 1333 | CGATTGAGCTTGGACCTGAGTTCAAGAAGGTTGCCAATTTCTGCTCGTTCAAGTCCA    | 1392 |
| Qy | 540  | lyGlnValSerValGluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnP | 560  |
| Db | 1393 | TCCCCCAGCATCGTCGAGCTTGAC-----AGCTTGAAGGTTTCTG                | 1431 |
| Qy | 560  | roGluValGlnTyrThrSerAsnTyrAlaLysSerAlaAsnValAspPheThr        | 577  |
| Db | 1432 | GTGATGTCGTGGTTTGGTTCTGGAATTACGCTCAAGGGCAAGGTGACAATCACC       | 1484 |

RESULT 13  
AK048546  
LOCUS  
DEFINITION  
Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched  
library, clone: C130072F01 product: COLLAGEN ALPHA 1(III) CHAIN  
PRECUSOR, full insert sequence.  
AK048546  
ACCESSION  
AK048546.1 GI:26339397  
VERSION

AK048546 4782 bp mRNA linear HTC 05-DEC-2002  
Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched  
library, clone: C130072F01 product: COLLAGEN ALPHA 1(III) CHAIN  
PRECUSOR, full insert sequence.  
AK048546  
ACCESSION  
AK048546.1 GI:26339397  
VERSION

KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE  
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Sasaki,N., Carninci,P.,  
TITLE Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
JOURNAL High-efficiency full-length cDNA cloning  
MEDLINE Meth. Enzymol. 303, 19-44 (1999)  
PUBMED 99279253  
PUBMED 10349636

REFERENCE  
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Sasaki,N., Carninci,P.,  
TITLE Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
JOURNAL Normalization and subtraction of cap-trapper-selected cDNAs to  
MEDLINE prepare full-length cDNA libraries for rapid discovery of new genes  
PUBMED Genome Res. 10 (10), 1617-1630 (2000)  
PUBMED 20499374  
PUBMED 11042159

REFERENCE  
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,  
TITLE Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,  
JOURNAL Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,  
MEDLINE Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,  
PUBMED Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,  
PUBMED Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,  
PUBMED Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
PUBMED RIKEN integrated sequence analysis (RISA) system--384-format  
PUBMED sequencing pipeline with 384 multicapillary sequencer  
PUBMED Genome Res. 10 (11), 1757-1771 (2000)  
PUBMED 20530913  
PUBMED 11076861

REFERENCE  
AUTHORS Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,  
TITLE Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,  
JOURNAL Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,  
MEDLINE Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,  
PUBMED Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,  
PUBMED Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,  
PUBMED Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,  
PUBMED Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,  
PUBMED Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,  
PUBMED Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,  
PUBMED Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,  
PUBMED Fletcher,C., Fujita,M., Gariboldi,M., Giustincich,S., Hill,D.,  
PUBMED Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,  
PUBMED Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,  
PUBMED Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,  
PUBMED Sato,K., Schonbach,C., Seiya,T., Shibata,Y., Storch,K.F., Suzuki,H.,  
PUBMED Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,  
PUBMED Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.  
PUBMED and Hayashizaki,Y.  
PUBMED Functional annotation of a full-length mouse cDNA collection  
PUBMED Nature 409 (6821), 685-690 (2001)  
PUBMED 21085660  
PUBMED 11217851

REFERENCE  
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research  
TITLE Group Phase I & II Team.  
JOURNAL Analysis of the mouse transcriptome based on functional annotation  
REFERENCE of 60,770 full-length cDNAs  
AUTHORS Nature 420, 563-573 (2002)  
PUBMED 6 (bases 1 to 4782)  
PUBMED Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,  
PUBMED Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,  
PUBMED Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,  
PUBMED Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,  
PUBMED Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,  
PUBMED Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,  
PUBMED Kamamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,  
PUBMED Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,  
PUBMED Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,  
PUBMED Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,

Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,  
Muramatsu,M. and Hayashizaki,Y.  
Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,  
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,  
Fax:81-45-503-9216)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL:http://genome.gsc.riken.go.jp/  
URL:http://fantom.gsc.riken.go.jp/  
Location/Qualifiers  
1. .4782  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:C130072F01"  
/db\_xref="taxon:10090"  
/clone="C130072F01"  
/tissue type="head"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="16 days embryo"  
119. .4513  
/note="unnamed protein product; COLLAGEN ALPHA 1(III)  
CHAIN PRECURSOR (SWISSPROT|P08121, evidence: FASTY,  
99.9%ID, 100%length, match=4392)  
putative"  
/codon\_start=1  
/protein\_id="BAC33370.1"  
/db\_xref="GI:26339398"  
/translation="MMSFVQSGTWFLTLHPTLILAQQSNVDELGCSHLGQSYESRD  
VWKEPQICVDSGSLCDDIICDEEPLDCNPEIPFGECCAICQPSTPAVLPDG  
HGQPGKDPDPPGIPGRNGDPLGQPLGPPSPGICESTGTGGQNYSPQFDSYD  
VKSQVGMGGYPGPAGPPGPPGSSGHPGSPGYQGPGEPPQAAGPPGPPG  
ALGPAGPKDGESRPGRGERGLPGPPGIGKPGAGMPFGMKGHRGDFDGRNGEKE  
TGAPGLKGNGLPGDNGAPGMPGRGAPGERGRLPGAAGARGKDGARGSDGQGP  
GPPGTAGFPSPGAKGEVGPAGSPGNSGSPQGRGEPGQGHAGACQPPPGNNGSPG  
GKGMGPAGIPGAPGLIGAPGPGAGTNGIPGTRGSPGPKNGAKGEPGARGERGE  
AGSPGIPGPKGEDKDGSPGEPGANGLPGAAGERGSPGFRGPAGNGIPGEKGPGER  
GGPGAGPRGVAGEPRDGTGPGGIRGMPGSPGGPNDGKPGPPSGGESRPGPPG  
PSGPRGPGVMGPPGPKNDGAPKNGERGPGGPGPLGPGAGKNGETGPGGPTGP  
AGDKGDSPPGPGQLOGIPGTGPPGKNGKPGGPKGEVAGAPGKGDSDGAPGER  
GPPGTAGIPGATGAGPPGPEGKGPAGPPGPPGASGSPGLQGMFGERGGSPGPKG  
EKGEPPGAGADGVPKGDPGRGPAGTIGPPGAPGDKGEGSSGLPGIAGPRGGPGE  
RGEHPPGPAGFPGAGQNGEPGAKGERGAPGKEGPPGPAGTSSGPGAGPPGQ  
GVKGERSPGPGTAGFPGGRGLPGPPGNNGNPPTGSPGAPGKDGPPGPNAGNSGSPG  
NPGIAGPKGDAGPGEGKPGPAQPPSPGPLGIAGLTGARGLAGPMPGPRGTGPG  
QGIKGESGKPGASCHNGERPPGQGLPGQPGTAGEPGRDGNPGSDGQPGRDGSPGK  
GDRGENSPGAPGAPGHPGPPGVPVPSGSGSDRGETPAGPSGAPGAPGARGAPGPG  
PRDKGTGTERGSGNGIKHRGFPGNPFPSPGSPAAGHQGAIGSPGAPGRGPVPHGP  
PGDKGTSGHPGPIGPPGPRGNRGERSESGSPHPGPPGPPGPPGCCGGAIA  
GVGKESGGSFPYGGDDPMDFKINTEEIMSSLSKSVNGQIESLISDGSRKNPARNCRD  
LKFCHEPLKSGEYWVDPNQCKMDAIVFCNMETGETCINASPMTVPRKHWTDGAE  
KKYVWFESMNGGFQFSYPPDLPEVDVQLAFLRLSSRASQNTYTHCKNSIAYMD  
OASGNVKKSLKLMGSEGEFKAEGNSKFTYTVLEDGCTKHTGWSKTVFEYQTRKMR  
LPIIDIAPYDIGGPDQEFVDIGPVCFI"  
4761. .4766  
/note="putative"  
4782  
/note="putative"  
BASE COUNT 1120 a 1216 c 1427 g 1019 t  
ORIGIN  
polyA\_signal  
polyA\_site  
Alignment Scores:  
Pred. No.: 73.9 Length: 4782  
Score: 113.00 Matches: 101









Qy 251 ValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGlnMetLeuArgThrGly 270  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 758 GTGACTCCTTTTCACATACAAAGACAAACAGAAAGCAAGAACGACGACCTGTACCAAG 817  
Qy 271 AsnAsnPheThrPheSerTyrThr---PheGluGluValPropheHisSerSerTyrAla 289  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 818 GGTGGGTTACATTTGGTAAAGTGGGCTCCTCTCTTTGCCATCTTCTCTCAATGTTGTT 877  
Qy 290 HisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyrLeuTyrTyrLeu 309  
: : : : : : : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 878 TTGGGAAGGACAGAGGAGAAACAGCAAGAACCTGTACTTCTACTTCTCTGGTCTTTGGG 937  
Qy 310 AsnArgThrGlnAsnGln-----SerGlySerAla 319  
: : : : : : : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 938 AAGAAAGCCGACAGTGAGGAGCCAAAGTGTCAAGCCAGTATTTTCCTTTGGGAATTCAGAG 997  
Qy 320 GlnAsnLysAsp-----LeuLeuPheSerArgGlySerProAlaGly 333  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 998 CAAACCAAGATGAGAGCTCCAAGCCGACATTCAGCTTCAGTGTGGCAAAACCATCTGGG 1057  
Qy 334 MetSerValGlnProLysAsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSer 353  
: : : : : : : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 1058 AAGGAG-----TCTGAGCAGCTAGCCAAG 1081  
Qy 354 LysThrLysThrAspAsnAsnAsnSerAsnPheThrTrp---ThrGlyAlaSerLysTyr 372  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 1082 GCTACTTTTGCACTTTGGAAATCAAACCAATACAACAACGTGATCAAGGTGCAGCAAAAGCCA 1141  
Qy 373 AsnLeuAsnGlyArgGluSerIleIleAsnProGlyThrAlaMetAlaSerHisLysAsp 392  
: : : : : : : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 1142 GTTTTAGCTTCTTGAACAGCAGTTCCTCTAGTTCAAGTGCACCGCCACT----- 1192  
Qy 393 AspGluAspLysPhePheProMetSerGlyValMetIlePheGlyLysGluSerAlaGly 412  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 1193 -----TCATCTAGTGGTGGCATATTGGT----- 1216  
Qy 413 AlaSerAsnThrAlaLeuAspAsnValMetIleThrAspGluGluIleLysAlaThr 432  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 1217 ---AGTCCACCTCTTCCCTCCAAACCCACCTGTGGCAGCCTTTGTGTTGGACAGGCCAGC 1273  
Qy 433 AsnProValAlaThrGluArgPheGlyThrValAlaValAsnPheGlnSerSerSer--- 451  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 1274 AATCCTGTCAAGCAGCTCTGCCCTTCGGTAACGCTGCTGAATCCAGTACATCTCAGTCTTG 1333  
Qy 452 -----ThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeuPro 466  
: : : : : : : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 1334 CTATTTCTCAAGAGAGCGAGCGAGCCAGCCACCGTCCAGCACTGCCCGAGCTGCCCTCTCCC 1393  
Qy 467 GlyMetValTrpGlnAspArgAsp-----ValTyrLeuGln 478  
: : : : : : : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 1394 TT-TGATTTGG---GACAGGAGCCAGCAGTAATAGTGTCTTCTGGGTTTCACTTTTGG 1449  
Qy 479 GlyProIleTrpAlaLysIle-ProHisThrAsp-----GlyHisPheHi 493  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 1450 GGCTACAACCAAGTCAAGCTCTTCAGGATCCTCCTTTGTATTTGGCACTGGACATTCAGC 1509  
Qy 493 sProSer-----ProLeuMetGly-----G1 500  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 1510 ACCATCTGCCAGTCCAGCATTTGGTGCTAACCAAGCTCCAACATTTGGACAAAGTCAAGG 1569  
Qy 500 yPheGlyLeuLysAsnProProGlnIleLeuIleLysAsnThr----- 515  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 1570 TGCCAGCCAGCCTAAACCGGCCAAGCTTTGGCTCTATCTCGTCTCAACAGCATTAATTTTC 1629  
Qy 516 -----ProValPro-----AlaAsnPr 521  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 1630 TGCTGGTTCTCAGCCTGTACCACCACCTATATTTGGGACAGTGTCAAGCAGCCCAACC 1689  
Qy 521 oPro-----AlaGluPheSerAlaThrLysPheAlaSe 532  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 1690 TCCTGTGTTTGGACAGCAGCCTAGTCAGTCTGCATTTGGCTCTGGGACAGCTAATGCCAG 1749

Qy 532 rPheIleThrGlnTyrSerThrGlyGlnValSerValGluIleGluTrpGluLeuGlnLy 552  
| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 1750 TTCGGTTTTCCAATTTGGTAGCAGCAGCTACGAATTC----- 1786  
Qy 552 sGluAsnSerLysArgTrpAsnProGluValGlnTyrThr-----SerAs 567  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 1787 ----AACTTCACAAACAATAATCCATCAGGAGTGTTCACATTTGGTGCAAGTCCCAGCAC 1842  
Qy 567 nTyrAlaLysSerAlaAsn-----Va 574  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 1843 GCCTGCAGCCCTCAGCCCGAGCCCTCAGGTTCAGGGGTCTTTTCATTAGCCAGTCTCCAGC 1902  
Qy 574 lAspPheThrValAspAsnAsnGly 582  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
Db 1903 ATCATTTACAGTGGGGTCAAATGGG 1927

Search completed: February 15, 2004, 01:28:59  
Job time : 3387.15 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 14, 2004, 15:19:20 ; Search time 5428 Seconds  
(without alignments)  
4024.639 Million cell updates/sec

Title: US-09-807-802A-17  
Perfect score: 2906  
Sequence: 1 MASGGGAPMADNNEGADGVG.....NNGLYTEPRPIGTRYLTRPL 534

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09807802/runat\_11022004\_175607\_15929/app.query.fasta\_1.2389  
-DB=GenEmbl -QFMT=fastap -SUPFIX=ige -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09807802 @CGN 1 1 12216 @runat\_11022004\_175607\_15929 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:\*

1: gb\_ba:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vi:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

| Result No. | Score  | Query Match % | Length | DB | ID       | Description        |
|------------|--------|---------------|--------|----|----------|--------------------|
| 1          | 2906   | 100.0         | 4718   | 14 | AF063497 | AF063497 Adeno-888 |
| 2          | 2884   | 99.2          | 4683   | 14 | AF028704 | AF028704 Adeno-888 |
| 3          | 2530   | 87.1          | 4722   | 14 | AF028705 | AF028705 Adeno-888 |
| 4          | 2528   | 87.0          | 4726   | 14 | AVU48704 | U48704 Adeno-888oc |
| 5          | 2518   | 86.6          | 8179   | 6  | AR264580 | AR264580 Sequence  |
| 6          | 2481.5 | 85.4          | 4072   | 6  | AR222046 | AR222046 Sequence  |
| 7          | 2481.5 | 85.4          | 4072   | 6  | AX205074 | AX205074 Sequence  |
| 8          | 2481.5 | 85.4          | 4679   | 6  | AX282480 | AX282480 Sequence  |
| 9          | 2481.5 | 85.4          | 4679   | 14 | AF043303 | AF043303 Adeno-888 |
| 10         | 2481.5 | 85.4          | 7327   | 12 | AF369963 | AF369963 Cloning v |
| 11         | 2481.5 | 85.4          | 7557   | 6  | AR222045 | AR222045 Sequence  |
| 12         | 2481.5 | 85.4          | 7557   | 6  | AX205073 | AX205073 Sequence  |
| 13         | 2481.5 | 85.4          | 8698   | 6  | AR222044 | AR222044 Sequence  |
| 14         | 2481.5 | 85.4          | 8698   | 6  | AX205072 | AX205072 Sequence  |
| 15         | 2457.5 | 84.6          | 2208   | 14 | AY243021 | AY243021 Non-human |
| 16         | 2439.5 | 83.9          | 4680   | 6  | AR028767 | AR028767 Sequence  |
| 17         | 2439.5 | 83.9          | 4680   | 6  | I62303   | I62303 Sequence 1  |
| 18         | 2439   | 83.9          | 2187   | 14 | AY243013 | AY243013 Non-human |
| 19         | 2437.5 | 83.9          | 2196   | 14 | AY243010 | AY243010 Non-human |
| 20         | 2436   | 83.8          | 2187   | 14 | AY243006 | AY243006 Non-human |
| 21         | 2436   | 83.8          | 2187   | 14 | AY243017 | AY243017 Non-human |
| 22         | 2431.5 | 83.7          | 2196   | 14 | AY243011 | AY243011 Non-human |
| 23         | 2430   | 83.6          | 4721   | 14 | AF513851 | AF513851 Adeno-888 |
| 24         | 2429.5 | 83.6          | 2217   | 14 | AY243015 | AY243015 Non-human |
| 25         | 2428.5 | 83.6          | 2217   | 14 | AY243009 | AY243009 Non-human |
| 26         | 2427.5 | 83.5          | 2217   | 14 | AY243007 | AY243007 Non-human |
| 27         | 2425   | 83.4          | 2187   | 14 | AY243016 | AY243016 Non-human |
| 28         | 2423.5 | 83.4          | 2202   | 14 | AY243012 | AY243012 Non-human |
| 29         | 2423.5 | 83.4          | 4675   | 6  | AX135805 | AX135805 Sequence  |
| 30         | 2423.5 | 83.4          | 4675   | 6  | AX286292 | AX286292 Sequence  |
| 31         | 2423.5 | 83.4          | 4675   | 6  | BD094552 | BD094552 Method of |
| 32         | 2423.5 | 83.4          | 4675   | 14 | AA2CG    | J01901 Adeno-888oc |
| 33         | 2422.5 | 83.4          | 2202   | 14 | AY243004 | AY243004 Non-human |
| 34         | 2422   | 83.3          | 2187   | 14 | AY243019 | AY243019 Non-human |
| 35         | 2421.5 | 83.3          | 2202   | 14 | AY243014 | AY243014 Non-human |
| 36         | 2421.5 | 83.3          | 2217   | 14 | AY243023 | AY243023 Non-human |
| 37         | 2419.5 | 83.3          | 2190   | 14 | AY242998 | AY242998 Non-human |
| 38         | 2418.5 | 83.2          | 2190   | 14 | AY242999 | AY242999 Non-human |
| 39         | 2418   | 83.2          | 2187   | 14 | AY243018 | AY243018 Non-human |
| 40         | 2417   | 83.2          | 2214   | 14 | AY243020 | AY243020 Non-human |
| 41         | 2411.5 | 83.0          | 2190   | 14 | AY243000 | AY243000 Non-human |
| 42         | 2411.5 | 83.0          | 2217   | 14 | AY243022 | AY243022 Non-human |
| 43         | 2410   | 82.9          | 2208   | 14 | AY243008 | AY243008 Non-human |
| 44         | 2410   | 82.9          | 2211   | 14 | AY242997 | AY242997 Non-human |
| 45         | 2390   | 82.2          | 1932   | 14 | AY243024 | AY243024 Non-human |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ALIGNMENTS

AF063497 AF063497 4718 bp DNA linear VRL 27-APR-1999  
LOCUS Adeno-associated virus 1, complete genome.  
DEFINITION AF063497  
ACCESSION AF063497.1 GI:4689096  
VERSION  
KEYWORDS  
SOURCE Adeno-associated virus 1  
ORGANISM Adeno-associated virus 1  
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.  
REFERENCE 1 (bases 1 to 4718)  
AUTHORS Xiao,W., Chirmule,N., Berta,S.C., McCullough,B., Gao,G. and Wilson,J.M.  
TITLE Gene therapy vectors based on adeno-associated virus type 1  
JOURNAL J. Virol. 73 (5), 3994-4003 (1999)  
MEDLINE 99214338  
PUBMED 10196295  
REFERENCE 2 (bases 1 to 4718)  
AUTHORS Xiao,W. and Wilson,J.M.  
TITLE Direct Submission  
JOURNAL Submitted (05-MAY-1998) IHGT, University of Pennsylvania, 3601 Spruce Street, Philadelphia, PA 19104, USA  
FEATURES  
source Location/Qualifiers  
1..4718  
/organism="Adeno-associated virus 1"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:85106"  
/note="AAV1"  
335..2206  
/codon\_start=1  
/product="nonstructural protein"  
/protein\_id="AAD27758.1"  
/db\_xref="GI:4689098"  
/translation="MPGFYEIVIKVPSDLDEHLPGISDSFVSWAEKEWELPPDSMD  
LNLI EQAPLTVAEKLQDFLVQRRVSKAPEALFFVQFEKGESYFHLHILVETTVKVS  
MVLGRFLSQIRDKLVQTIYRGIEPTLPNPFVTKTRNGAGGKNVVDCEYIPNYLLPK  
TQPELQAWTNMEEYISACLNLAERKRLVACHLTHVSTQEQNKENLNPNSDAPVIRS  
KTSARYMELVGLVDRGITSEKQWIQEDQASYISFNAASNSRSQIKAAALDNAGKIMAL  
TKSAPDYLVGPPADIKTNRIYRILELNGYEPAYAGSVFLGWAQKRFGRNTIWLFG  
PATTKGTNIAEATAHAVPFYGCVNWTNENPFENDCVDKMWIWEEGKMTAKVWESAKA  
ILGSKVRVDQKSSAQIDPTPVI VTSNTNMCAVIDGNSTTFEHQOPLQDRMFKFEL  
TRRLEHDFGVTQGEVKFFRWAQDHVTEVAHEFYVRKGGANKRPAPDDADKSEPKRA  
CPSVADPSTSDAEGAPVDFADRYQNKCSRHAGMLQMLFPCKTCERNQNFNICFTHGT  
RDCSECFPGVSESQPVVRKRTYRKLCIAIHILGRAPEIACSACDLVNVVDLDDCVSEQ"  
2223..4433  
/codon\_start=1  
/product="capsid protein"  
/protein\_id="AAD27757.1"  
/db\_xref="GI:4689097"  
/translation="MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQKQDDGRG  
LVLPGYKYLGPFGNLGDKGEFVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADAEOF  
RLQEDTSFGGNLGRAVFOAKRVLPLEPLGLEGAKTAPGKRPFVQSPQEPDSSSGIG  
KTGQPAKKRLNFGQTGDSVPDQPLGEPPATPAVGPPTMASSGGGAPMADNNEGA  
DVGNGASGNWHCHDSFWLGDRTVITSTRTWALPTYNHLYKQISSASTGASNDNHYFGY  
STPWGYFDENRFHCHFSPRDWQRLINNNWGFPRKLNFKLENIQVKEVTNDGVTTIA  
NNLTSTVQVFSDSYQLPYVLGSAHQGLPFPADVFMIPOYGYLTNLNGSQAVGRSS  
FYCLEYFPSQMLRTGNFTFSYTFEEVPPHSSYAHSQSLDRLMNPLIDQYLYLNRTO  
NQSOSAONKDLLFSRGS PAGMSVQPKNWLPGQRVSKTKTDNNNSNFTWTGASK  
YNLNGRESIINPGTAMASHKDDKDFPMSGVNIFGKESAGASNTALDNVMTDEEEI  
KATNPVATERFGTVAVNFIQSSSTD PATGDVHAMGALPGMWQDRDYLQGP IAWAKIPH  
TDGHFHPSPLMGGFGLKNPPQIILIKNTVPANPPAEFSATKFPASFTQYSTGQVSVE  
IEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPICTRYLTRL"

BASE COUNT 1121 a 1393 c 1273 g 931 t

ORIGIN

Alignment Scores:  
Pred. No.: 3.39e-226 Length: 4718  
Score: 2906.00 Matches: 534  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0  
US-09-807-802A-17 (1-534) x AF063497 (1-4718)

1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20  
2829 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCCGACGGAGTGGGT 2888  
21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40  
2889 AATGCCTCAGGAATTTGGCATTGGCATTTCCACATGGCTGGGCGACAGAGTCATCACCACC 2948  
41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
2949 AGCACCCGCACCTGGGCCTTGGCCACCTACAATAACCACCTCTACAAGCAAAATCTCCAGT 3008  
61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTyrGlyTyr 80  
3009 GCTTCAACGGGGGCCAGCAACGACCAACCACTACTTTCGGCTACAGCACCCCTGGGGGTAT 3068  
81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100  
3069 TTTGATTCAACAGATTCCACTGCGCACTTTTTCACCACGTCGTCGTCGTCGTCGTCGTCG 3128  
101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120  
3129 AACAAATTTGGGGATTCCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAA 3188  
121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140  
3189 GAGGTCAAGACGAATGATGGCGTCACAACCACTCGCTAATAACCTTACCAGCACGGTTCAA 3248  
141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160  
3249 GTCTTCTCGGACTCGGAGTACCAGTTCGCTACGTCCTCGGCTCTGCGCACCGGGCTGC 3308  
161 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
3309 CTCCTTCGTTCCCGCGGACGTCGTTTCATGATTCGCAATACGCTACCTGACGCTCAAC 3368  
181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
3369 AATGGCAGCCAAGCCGTGGGACGTTTCATCTTTTACTGCTTGGAAATATTTCCCTTCTCAG 3428  
201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluValProPheHis 220  
3429 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 3488  
221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
3489 AGCAGCTACGCGCACAGCCAGAGCTGGACCGGCTGATGATCTCTCATCGACCAATAC 3548  
241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260  
3549 CTGTATTACTCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCTGAGGACCTTGCTG 3608  
261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280  
3609 TTTAGCCGTGGGTCTCCAGCTGGCATGTCGTTTCAGCCCCAAAACTGGCTACCTGGACCC 3668  
281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
3669 TGTATCGGAGCAGCGCGTTTCTAAAAACAAAACAGACAACAACACAGCAATTTTACC 3728  
301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320  
3729 TGGACTGGTGTCTCAAAATATAACCTCAATGGCGTGAATCCATCATCAACCTGGCACT 3788  
321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340  
3789 GCTATGGCCTCACACAAAGACGACGAGACAAGATTCTTTCCCATGAGCGGTGTCATGATT 3848  
341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360  
3849 TTTTGAAAAGAGAGCGCGGAGCTTCAAAACACTGCATTGGACAATGTTCATGATTACAGAC 3908

|  |      |   |      |
|--|------|---|------|
| QY   | 361  | GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal  | 380  |
| Db   | 3909 | GAAGAGGAAATTAAAGCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC   | 3968 |
| QY   | 381  | AsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu     | 400  |
| Db   | 3969 | AATTTCCAGAGCAGCAGCAGACCCCTGCGACCGAGATGTGCATGCTATGGGAGCATTA    | 4028 |
| QY   | 401  | ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle  | 420  |
| Db   | 4029 | CCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCAGGGTCCCATTTGGGCCAAAATT  | 4088 |
| QY   | 421  | ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn  | 440  |
| Db   | 4089 | CCTCACACAGATGGACACTTTCACCCGTCTCCTCTTATGGCGGCTTTGGACTCAAGAAC   | 4148 |
| QY   | 441  | ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe  | 460  |
| Db   | 4149 | CGCCTCTCAGATCCTCATCAAAACACGCCCTGTCTCTGCGAATCCTCCGCGGAGTTT     | 4208 |
| QY   | 461  | SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu  | 480  |
| Db   | 4209 | TCAGCTACAAAGTTTGCTTCATTCAATCACCCCAATACTCCACAGGACAAAGTAGTGTGAA | 4268 |
| QY   | 481  | IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr  | 500  |
| Db   | 4269 | ATTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGAGTACACA    | 4328 |
| QY   | 501  | SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr  | 520  |
| Db   | 4329 | TCCAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAACAATGGACTTTATACT  | 4388 |
| QY   | 521  | GluProArgProIleGlyThrArgTyrLeuThrArgProLeu                    | 534  |
| Db   | 4389 | GAGCCTCGCCCATTTGGCACCGTTACCTTACCCGTCCTG 4430                  |      |
| RESULT 2   |      |   |      |
| AF028704   |      |   |      |
| LOCUS  |      |   |      |
| DEFINITION   |      |   |      |
| ACCESSION  |      |   |      |
| VERSION  |      |   |      |
| KEYWORDS   |      |   |      |
| SOURCE   |      |   |      |
| ORGANISM   |      |   |      |
| REFERENCE  |      |   |      |
| AUTHORS  |      |   |      |
| TITLE  |      |   |      |
| JOURNAL  |      |   |      |
| MEDLINE  |      |   |      |
| PUBMED   |      |   |      |
| REFERENCE  |      |   |      |
| AUTHORS  |      |   |      |
| TITLE  |      |   |      |
| JOURNAL  |      |   |      |
| FEATURES   |      |   |      |
| source   |      |   |      |
| 1. .4683   |      |   |      |
| /organism="Adeno-associated virus 6"                       |      |   |      |
| /mol_type="genomic DNA"                                    |      |   |      |
| /db_xref="taxon:68558"                                     |      |   |      |
| /note="AAV6"   |      |   |      |
| 320. .2191   |      |   |      |
| /codon_start=1   |      |   |      |
| /evidence=not experimental                                 |      |   |      |
| /product="nonstructural protein Rep78"                     |      |   |      |
| /protein_id="AAB95449.1"                                   |      |   |      |
| /db_xref="GI:2766606"                                      |      |   |      |
| /translation="MPGFYEIVIKVPSDLDEHLPGISDSFVNWVAEKEWELPPDSDMD |      |   |      |
| LNLI EQAPLTVAEKLQDFLVQWRRVSKAPEALFFVQFEKGESYFHLHLIVETTGKVS |      |   |      |
| MVLGRFLSQIRDKLVTIYRGIEPTLPNFWAVTKTRNGAGGGKNKVVDCEYIPNYLLPK |      |   |      |

|  |      |  |      |
|--|------|--|------|
| TQPELQWANTNMEEYISACLNLAERKRLVAHDLTHVTSQTEQNKENLNPNSDAPVIRS   |      |  |      |
| KTSARYMELVGLVDRGITSEKQWIOEQASYISFNAASNSRQIKAAOLDNAGKINAL     |      |  |      |
| PKTSAPDYLVGPAAPPADIKTNRIYRILELNGYDPAYGSVFLGWAQKRFGRNITWLFQ   |      |  |      |
| PATTKNTIAEAIAHAVPFYGCNVNTNENFPFNDCTVDKVIWEEGKMTAKVVESAKA     |      |  |      |
| ILGSKVRVDQCKSSAQIDPTPVIIVTSNTNMCVTDGNSTTFHQQPLQDRMFKEP       |      |  |      |
| TRRLEHDFGKVTKEVKEFFRWAQDHVTEVAHEFYVRKGANKRPAPDDADKSEPKRA     |      |  |      |
| CPSVADPSTDAEGAPVDPADRYQNKCSRHAGMLQMLFPCKTCERMNQNFNICFTHGT    |      |  |      |
| RDCSECFPGVSESQPVVRKTYRKLCAIHLLGRAPEIACSACDLVNVVDLDDCVSEQ"    |      |  |      |
| 2208. .4418  |      |  |      |
| /codon_start=1   |      |  |      |
| /evidence=not experimental                                   |      |  |      |
| /product="capsid protein vp1"                                |      |  |      |
| /protein_id="AAB95450.1"                                     |      |  |      |
| /db_xref="GI:2766607"  |      |  |      |
| /translation="MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRG    |      |  |      |
| LVLPGYKYLGPFGNGLDKGEPVNAADAALEHDKAYDQQLKAGDNPYLRYNHADAEPQE   |      |  |      |
| RLQEDTSFGNLRGRAVFOAKKRVLEPFGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIG   |      |  |      |
| KTGQQPAKKRLNFGQTGDSVPDPQLGEPPTAAVGPPTMASGGGAPMADNNEGA        |      |  |      |
| DGVGNASGNWHCDSTWLGDRVITTSRTTWALPTYNNHLYKQISSASTGASNDNHYFGY   |      |  |      |
| STPWGYDFDNRFCHFSPRDWQRLINNNWGRFRKRLNPKLFNIQVKEVTTNDGVTTIA    |      |  |      |
| NNLTSTVQVFSDEYQLPYVLGSAHQCLPPFPADVPMI PQYGYLTNLNGSQAVGRSS    |      |  |      |
| FYCLEYFPQMLRTGNFTFSYTFEDVPFHSSYAHSSQSLDRLMNLPLIDQYLYLNRTQ    |      |  |      |
| NQSQAQNKDLLFSRGSAGMSVQPKWLPGPCYRQQRVSKTKTDNNNSNFTWTGASK      |      |  |      |
| YNLNGRESIIINPGTAMASHKDDKDKFFPMGSGVMI FGKESAGASNTALDNVMTDSEEI |      |  |      |
| KATNPVATERFQVAVNLQSSSTDPTATGDVHVMPALPGMWQDRDVLQGYIWKAPIH     |      |  |      |
| TDGHFHPSPLMGGFGLKHPPIILIKNTVPVANPPALPFSATKPFASFITQYSTQVSV    |      |  |      |
| IEWELQKENSKRWNPEVQVTSNYAKSANVDFTVDDNNGLYTEPRPIGTRYLTRPL"     |      |  |      |
| BASE COUNT 1114 a 1363 c 1277 g 929 t                        |      |  |      |
| ORIGIN   |      |  |      |
| Alignment Scores:  |      |  |      |
| Pred. No.: 2.06e-224 Length: 4683                            |      |  |      |
| Score: 2884.00 Matches: 529                                  |      |  |      |
| Percent Similarity: 99.63% Conservative: 3                   |      |  |      |
| Best Local Similarity: 99.06% Mismatches: 2                  |      |  |      |
| Query Match: 99.24% Indels: 0                                |      |  |      |
| DB: 14 Gaps: 0   |      |  |      |
| US-09-807-802A-17 (1-534) x AF028704 (1-4683)                |      |  |      |
| QY   | 1    | MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly    | 20   |
| Db   | 2814 | ATGGCTTCAGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGAGTGGGT       | 2873 |
| QY   | 21   | AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr | 40   |
| Db   | 2874 | AATGCCTCAGAAATTGGCATTGCGATTCCACATGGCTGGCGGACAGATCATCACC      | 2933 |
| QY   | 41   | SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer | 60   |
| Db   | 2934 | AGCACCCGAACATGGGCCTTGCCCACTATATAACAACCACTCTCTACAAGCAATCT     | 2993 |
| QY   | 61   | AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr | 80   |
| Db   | 2994 | GCTTCAACGGGGCCAGCAACGACAACCACTACTTCGGCTACAGCACCCCTGGGGT      | 3053 |
| QY   | 81   | PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn | 100  |
| Db   | 3054 | TTTGATTTCAACAGATTCCACTGCCATTCTCACCACGTGACTGGCAGCGACTCAT      | 3113 |
| QY   | 101  | AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys | 120  |
| Db   | 3114 | AACAATTGGGGATTCCGGCCCCAAGAGACTCAACTTCAAGCTCTTCAACATCCA       | 3173 |
| QY   | 121  | GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln | 140  |
| Db   | 3174 | GAGGTACGACGAATGATGGCGTCACGACCATCGCTAATAACCTTACCAGCAGG        | 3233 |
| QY   | 141  | ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys | 160  |
| Db   | 3234 | GTCTTCTCGGACTCGGAGTACCAGTTGCCGTACGTCTCTCGGCTCTGCGCAC         | 3293 |
| QY   | 161  | LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn | 180  |



Db 3294 CTCCCTCGTTCGGCGGACGTTGTCATGATTCGCGAGTACGGTACCTAACGCTCAAC 3353  
Qy 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
Db 3354 AATGGCAGCCAGGACGAGTGGGACGGTCATCCTTTTACTGCGTGAATATTTCCCATCGCAG 3413  
Qy 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis 220  
Db 3414 ATGCTGAGAACGGGCAATAACTTTACCTTCAGCTACACCTTCGAGGACGTGCCTTTCCAC 3473  
Qy 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
Db 3474 AGCAGCTACCGGCACAGCCAGAGCCTGGACCGGTGATGTAATCCTCTCATCGACCCAGTAC 3533  
Qy 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260  
Db 3534 CTGTATTACTGAACAGAACTCAGAACTCAGTCCGGAAGTGCCCAAAACAAAGGACTTGCTG 3593  
Qy 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280  
Db 3594 TTTAGCCGGGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACTGGTACCTACCTGGACCC 3653  
Qy 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
Db 3654 TGTTACCGGACGACGCGGTTTCTAAAAACAAAACAGACAAACAAACAGCAACTTTTACC 3713  
Qy 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320  
Db 3714 TGGACTGGTGCTTCAAAATATAACCTTAATGGCGGTGAATCTATATAATCAACCCCTGGCACT 3773  
Qy 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340  
Db 3774 GCTATGGCCTCACACAAAGACGACAAAGACAAGTTCTTTCCCATGAGCGGTGTCATGATT 3833  
Qy 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360  
Db 3834 TTTGGAAAGGAGAGCGCCGGAGCTTCAAACACTGCATTTGGACAATGTCTATGATCACAGAC 3893  
Qy 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380  
Db 3894 GAAGAGGAATCAAAGCCACTAACCCCGTGGCCACCGAAAGATTGGGACTGTGGCAGTC 3953  
Qy 381 AsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400  
Db 3954 AATCTCCAGAGCAGCAGACAGACCCTGCGACCGGAGATGTGCATGTTATGGAGCCTTA 4013  
Qy 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420  
Db 4014 CCTGGAATGGTGTGGCAAGACAGACGATATACCTGCAGGGTCTCTATTTGGGCCAAAATT 4073  
Qy 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440  
Db 4074 CCTCACACGGATGGACACTTTCACCGCTCTCCTCTCATGGCGGCTTTGGACTTAAGCAC 4133  
Qy 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460  
Db 4134 CCGCCTCCTCAGATCCTCATCAAAACACCGCTGTTCTCGGAATCCTCCGGCAGAGTTT 4193  
Qy 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480  
Db 4194 TCGGCTACAAAGTTTGCTTCATTCACCCAGTATTCACAGGACAAGTGAGCGTGGAG 4253  
Qy 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 500  
Db 4254 ATTGAATGGGAGCTGCAGAAAGAAACAGCAACCGCTGGAATCCCGAAGTGCAGTATACA 4313  
Qy 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520  
Db 4314 TCTAACTATGCAAAATCTGCCAACGTTGATTTCACTGTGGACAACAATGGACTTTTATACT 4373  
Qy 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534

Db 4374 GAGCCTCGCCCCCATTTGGCACCCGTTTACCTCACCCGTTCCCTG 4415  
RESULT 3  
AF028705 4722 bp DNA linear VRL 12-JAN-1998  
LOCUS Adeno-associated virus 3B, complete genome.  
DEFINITION AF028705  
ACCESSION AF028705.1 GI:2766608  
KEYWORDS Adeno-associated virus 3B  
SOURCE Adeno-associated virus 3B  
ORGANISM Adeno-associated virus 3B  
REFERENCE 1 (bases 1 to 4722)  
AUTHORS Rutledge,E.A., Halbert,C.L. and Russell,D.W.  
TITLE Infectious clones and vectors derived from adeno-associated virus (AAV) serotypes other than AAV type 2  
J. Virol. 72 (1), 309-319 (1998)  
MEDLINE 98080418  
PUBMED 9420229  
REFERENCE 2 (bases 1 to 4722)  
AUTHORS Rutledge,E.A. and Russell,D.W.  
TITLE Direct Submission  
JOURNAL Submitted (25-SEP-1997) Hematology, Univ of Washington, Box 357720, Seattle, WA 98195, USA  
FEATURES  
source  
1. .4722 Location/Qualifiers  
/organism="Adeno-associated virus 3B"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:68742"  
/note="AAV3B"  
317. .2191  
/codon\_start=1  
/product="nonstructural protein Rep78"  
/protein\_id="AAB95451.1"  
/db\_xref="GI:2766609"  
/translation="MPGFYEIVLKVPSDLDEHLPGISNFSVNWVAEKBWELPPDSMD  
PNLIEQAPLTVAEKLQREFLVWRVRSKAPALFFVQFEKGETYFHLVLIETIGVKS  
MVVGRYVSOIKEKLVTIRYRGVEPQLPNWFVTKTRNGAGGNGKVVDCCYIPNYLLPK  
TOPELQAWTNMDQYLSACLNLAEKRLVAQHLTHVTSQTQEQNKENQNSDAPVIRS  
KTSARYMELVGLVDRGITSEKOWIOEQOASYISFNAASNSRSQIKALDNASKIMSL  
TKTAPDYLGSNPPEDITKNRIYQILENGYDPQYAAVSFLGWAQKFGKNTIWLFG  
PATTKTNTIABAIHAHVPFYGCNVNTNENFPFNDCVDKXVIWEEGQMTAKVVESEKA  
ILGGSKVRVDQCKSSAQIEPTPIVTSTNMCAVIDGNSTTTEHQQLQDRMFKFEL  
TRRLDHDGKVTQEVKDFFRWASDHVTDVAHEFYVRKGAKKRPAASNDADVSEPKRQ  
CTSLAQPTTSDAEPADYADRYQNKCSRHVGMNMLFPCKTCERMNQISNVCPTHGQR  
DCGECFPGMSESQPVSVVKKTYQKLCPIHHILGRAPEIACSDLANVDLDDCVSEQ  
"  
2208. .4418  
/codon\_start=1  
/product="capsid protein VP1"  
/protein\_id="AAB95452.1"  
/db\_xref="GI:2766610"  
/translation="MAADGYLPDWLENLSEIREWALKPGVPQPKANOQHQDNRRRG  
LVLPGYKYLGPNGLDKGPEVNEADAALAEHDKAYDQQLKAGDNPYLYKNHADAEFQE  
RLQEDTSFGGNLGRAVFOAKRILEPLGLVEEAKTAPGKRPVDQSPQSPDSSSGVG  
KSGKQPARKRLNFGQTGDSEVPDPQPLGEPAPAPTSLGSNTMASGGGAPMADNNEGA  
DGVNSSGNTHCDSQWLGDREVITTTSTRTWALPTYNHLYKQISSQSGASNDNHYFGYS  
TPWGYFDNRFHCHFPSPRDQRLINNNWGFPRKLSFKLFNIQVKEVTQNDGTTIAN  
NLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFADVFVMPQYGYLTNLNGSAVGRSSP  
YCLEYFSPQMLRTGNNFQSYTFEDVPFHSSYAHSSQLDRLMNPIDQYLYLNRTQG  
TTSGTTNQSRLLFSQAGPQSMQSLQARNWLPQCYRQRLSKTANDNNNSNFWTAASK  
YHLNGRDSLVPNGPAMASHKDDEKFFPMHGNLI FGEGTTASNAELDNVMTDEEEI  
RTTNPVATEQYGTVANNLQSSNTATPTRTVNDQCALPGMWQDRDYLQGIWAKIPH  
TDGHFHPSPLMGGFGLKHPPPQIMIKNTVPANPPTTFSPAKFASFITQYSTGQSVSE  
IEWELQKENSKRWNPEIQYTSYNKSVNVDFTVDTNGVSEPRPIGTRYLTRNL"

BASE COUNT 1204 a 1297 c 1243 g 978 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.26e-195 Length: 4722  
Score: 2530.00 Matches: 460  
Percent Similarity: 91.03% Conservative: 27  
Best Local Similarity: 85.98% Mismatches: 46

|   |        |  |      |
|---|--------|--|------|
| Query Match:                                  | 87.06% | Indels:  | 2    |
| DB:   | 14     | Gaps:  | 2    |
| US-09-807-802A-17 (1-534) x AF028705 (1-4722) |        |  |      |
| Qy  | 1      | MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly          | 20   |
| Db  | 2814   | ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAGGTGCCGATGGAGTGGGT        | 2873 |
| Qy  | 21     | AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr       | 40   |
| Db  | 2874   | AATTCCTCAGGAATTTGGCAATTGGCAATTCGAATTCCTCAATGGCTGGCGCAGAGTCATCACACC | 2933 |
| Qy  | 41     | SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer       | 60   |
| Db  | 2934   | AGCACCAGAACCTGGGCCCTGCCACTTACAAACCACTCTCTACAAGCAAATCTCCAGC         | 2993 |
| Qy  | 61     | AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr       | 80   |
| Db  | 2994   | CAATCA--GGAGCTTCAACGACAACCACTACTTTGGCTACAGCACCCCTTGGGGGTAT         | 3050 |
| Qy  | 81     | PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn       | 100  |
| Db  | 3051   | TTTGACTTTAACAGATTCCACTGCACTTCTCACCACGTGACTGGCAGCGACTCATTAAC        | 3110 |
| Qy  | 101    | AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys       | 120  |
| Db  | 3111   | AACAACCTGGGGATTCCGGCCCCAAGAACTCAGCTTCAAGCTTCTTCAACATCCAAAGTTAA     | 3170 |
| Qy  | 121    | GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln       | 140  |
| Db  | 3171   | GAGGTCAAGCAGAACGATGGCAGCAGCTATTGGCCAAATAACCTTACCAGCACGGTTCAA       | 3230 |
| Qy  | 141    | ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys       | 160  |
| Db  | 3231   | GTGTTTACGGACTCGGAGTATCAGCTCCCGTACGTGCTCGGGTCGGCGCACCAAGGCTGT       | 3290 |
| Qy  | 161    | LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn       | 180  |
| Db  | 3291   | CTCCCGCGGTTTCCAGCGGACGTCTTCATGGTCCCTCAGTATGGATACCTCACCCCTGAAC      | 3350 |
| Qy  | 181    | AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln       | 200  |
| Db  | 3351   | AACGGAGTCAAGCGGTGGGACGCTCATCTTTTACTGCTCGGAGTACTTCCCTTCGCAG         | 3410 |
| Qy  | 201    | MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis       | 220  |
| Db  | 3411   | ATGCTAAGGACTGGAAATAACTTCCAATTCAGCTATACCTTCGAGGATGTACCTTTTCAC       | 3470 |
| Qy  | 221    | SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr       | 240  |
| Db  | 3471   | AGCAGTACGCTCACAGCCAGAGTTTGGATCGCTTGATGATCTCTTATTGATCAGTAT          | 3530 |
| Qy  | 241    | LeuTyrTyrLeuAsnArgThrGln---AsnGlnSerGlySerAlaGlnAsnLysAspLeu       | 259  |
| Db  | 3531   | CTGTACTACTGAACAGACGACGCAAGCAAGCAACCTCTGGAACAACCAACCAATCACGGCTG     | 3590 |
| Qy  | 260    | LeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGly       | 279  |
| Db  | 3591   | CTTTTTCAGCAGGCTGGGCCTCAGTCTATGCTTTGTCAGGCCAGAAAATGGCTACCTGGG       | 3650 |
| Qy  | 280    | ProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPhe       | 299  |
| Db  | 3651   | CCCTGTACCGGCAACAGAGACTTTCAAGACTGCTAAACGACAAACAACAGTAACTTT          | 3710 |
| Qy  | 300    | ThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGly       | 319  |
| Db  | 3711   | CCTTGGACAGCGGCCAGCAAAATATCATCTCAATGGCCGCGACTCGCTGGTGAATCCAGGA      | 3770 |
| Qy  | 320    | ThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMet       | 339  |
| Db  | 3771   | CCAGCTATGGCCAGTCAACAGGACGATGAAGAAAAATTTTTCCTATGCACGGCAATCTA        | 3830 |

|            |  |   |                        |
|------------|--|---|------------------------|
| Qy         | 340  | IlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThr    | 359                    |
| Db         | 3831   | ATATTTTGGCAAGAAAGGGACAACCGCAAGTAACGCAGAAATTAGATAATTGTAATGATTACG | 3890                   |
| Qy         | 360  | AspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAla    | 379                    |
| Db         | 3891   | GATGAAGAAGAGATTGTCACCAATCCTGTGGCAACAGAGCAGTATGGAACCTGTGGCA      | 3950                   |
| Qy         | 380  | ValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAla    | 399                    |
| Db         | 3951   | AATAACTTGCAGAGCTCAAATACAGCTCCCAAGCTAGAACTGTCAATGATCAGGGGGCC     | 4010                   |
| Qy         | 400  | LeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLys    | 419                    |
| Db         | 4011   | TTACTTGGCATGCTGTGGCAAGATCGTGACGTGTACCTTCAAGGACCTATCTGGGCAAAAG   | 4070                   |
| Qy         | 420  | IleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLys    | 439                    |
| Db         | 4071   | ATTCTTCACACGGATGGACACTTTTCATCTCTCTCTGATGGGAGGCTTTGGACTGAAA      | 4130                   |
| Qy         | 440  | AsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGlu       | 459                    |
| Db         | 4131   | CATCCGCCTCCTCAAATCATGATCAAAATACTCCGGTACCGGCAAAATCCTCCGAGACT     | 4190                   |
| Qy         | 460  | PheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerVal    | 479                    |
| Db         | 4191   | TTCAGCCCCGCCAAGTTTGCTTTCATTTTACTCAGTACTCCACTGGACAGGTCAGCGTG     | 4250                   |
| Qy         | 480  | GluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyr    | 499                    |
| Db         | 4251   | GAATTTGAGTGGGAGCTACAGAAAGAAAACAGCAAACTGTTGGAATCCAGAGATTCAGTAC   | 4310                   |
| Qy         | 500  | ThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyr    | 519                    |
| Db         | 4311   | ACTTCCAACCTACAACAAGTCTGTTAATGTGGACTTTACTGTAGACACTAATGGTGTATT    | 4370                   |
| Qy         | 520  | ThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu                   | 534                    |
| Db         | 4371   | AGTGAACCTCGCCCTATTGGAAACCCGGTATCTCACAGAAACTTG                   | 4415                   |
| RESULT 4   |  |   |                        |
| AVU48704   | 4726 bp  | DNA   | linear VRL 15-JUL-1996 |
| LOCUS      | Adeno-associated virus 3   | nonstructural protein and capsid protein                        |                        |
| DEFINITION | genes, complete cds, and complete genome.  |   |                        |
| ACCESSION  | U48704   |   |                        |
| VERSION    | U48704.1   | GI:1408467  |                        |
| KEYWORDS   | Adeno-associated virus 3   |   |                        |
| SOURCE     | Adeno-associated virus 3   |   |                        |
| ORGANISM   | Adeno-associated virus 3   |   |                        |
| REFERENCE  | Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.  |   |                        |
| AUTHORS    | 1 (bases 1 to 4726)  |   |                        |
| TITLE      | Muramatsu,S., Mizukami,H., Young,N.S. and Brown,K.E.   |   |                        |
| JOURNAL    | Nucleotide sequencing and generation of an infectious clone of   |   |                        |
| MEDLINE    | adeno-associated virus 3   |   |                        |
| PUBMED     | Virology 221 (1), 208-217 (1996)   |   |                        |
| REFERENCE  | 8661429  |   |                        |
| AUTHORS    | 2 (bases 1 to 4726)  |   |                        |
| TITLE      | Muramatsu,S. and Brown,K.E.  |   |                        |
| JOURNAL    | Direct Submission  |   |                        |
| FEATURES   | Submitted (06-FEB-1996) Shin-ichi Muramatsu, Hematology Branch, NHLBI/NIH, 9000 Rockville, MD 20892, USA |   |                        |
| source     | Location/Qualifiers  |   |                        |
|            | 1. .4726   |   |                        |
|            | /organism="Adeno-associated virus 3"   |   |                        |
|            | /mol_type="genomic DNA"  |   |                        |
|            | /strain="3H"   |   |                        |
|            | /db_xref="taxon:46350"   |   |                        |
|            | 318. .2192   |   |                        |
|            | /codon_start=1   |   |                        |
|            | /product="nonstructural protein"   |   |                        |
|            | /protein_id="AAC55048.1"   |   |                        |
| CDS        |  |   |                        |

/db\_xref="GI:1408468"  
/translation="MPGFYEIVLKVPSDLERLPGISNSFVNWVAEKEWVPPDSMD  
PNLIEQAPLTVAEKLQREFLVEWRRVSKAPEALFFVQPEKGETYFHLHLVLIETIGVKS  
MVGRYVSQIKEKLVTRIYRGVEPQLPNWFAVTKTRNGAGGKNKVDVCYIPNYLLPK  
TQPELOWANTMDQYLSACLNAERKRLVAQLHLYTQVSTQEQNKENQNPNSDAPVIRS  
KTSARYMELVGLWDRGITSEKQWQEDQASYISFNAASNSRSQIKAAALDASKIMSL  
TKTAPDYLVSNPPEITKRIYQILELNGYDPOVAASVFLGWAQKKFGKRNITWLF  
PATGKTNIATAEIAHAVPFYGCWNWTNENFPFNDVCDVMVIWEEGKMTAKVYESAKA  
ILGSKVRVDQCKSSAQIEPTPIVTSNTNMCAVDGNSTTFEHQOPLQDMFEFEL  
TRRLHDFGKVTQKVEHFOFFWASDHVTDVAHEFYVRKGGAKRPPASNDADVSEPKRE  
CTSLAQPTTSDAEPADYADRYQNKCSRHVGNMLFLPCKTCERMNQISNVCFTHGQR  
DCGECPPGMSESQPVSVVKKTYQKLCPIHHILGRAPEIACSACDLANVDLDDCVSEQ  
"

CDS  
2209..4419  
/codon\_start=1  
/product="capsid protein"  
/protein\_id="AAC55049.1"  
/db\_xref="GI:1408469"  
/translation="MAADGYLPDWLEDNLSEGIREWALKPGVPPQPKANQQHODNRRG  
LVLPGYKYLPGNGLDKGEVNEADAAALEHDKAYDQQLKAGDNPYLKYNHADAEFQE  
RLQEDTSFGGNLGRAVFAKRILEPLGLVEAAKTAQKKGAVDQSPQEPDSSGVG  
KSGQDPARKRLNFGQTGDSVESVDPQLGEPAPAPTSLGNTMASGGGAPWADNNEGA  
DSGVNSGNWHCDSQLGDRVITSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYS  
TPWGYPDFNRFHCFSPRQRLNNWGFPPKLSFKLFNIQVRGVTQNDGTTTIAN  
NLTSYVQVFTDSEYQLPYVLGSAHQGLPPFPADVFMVPOYGYLTLNNGSQAVERSSF  
YCLEYFPQSMLRTGNFQSYTFEDVPFHSYSAHSQSLDLRLMNLPLIDQYLYLNRKIQG  
TTSGETNQSRLLFSQAGPQSMLSQARNWLPGPCYRQORLSKTANDNNSNFFPWTAAAK  
YHLNGRDSLVPNGPAMASHKDDDEKFFPMHGNIIFGEGSTASNAELDNVMTDEEII  
RTTNVPATEOYGTVANLQSSNTAPTTGTNVHQALPGMWQDRDVLQGPWAKIPH  
TDGHPHPSPLMGFGGLKHPPOIMIKNTVPANPPTTFSPAKFASFIQYSTGVQSVVE  
IEWELOKENSKRWNPEIQYTSNTKNSVNVDFTVDINGVYSEPRPIGTRYLTRNL"

BASE COUNT 1200 a 1295 c 1251 g 980 t  
ORIGIN

Alignment Scores:

Pred. No.: 1 84e-195 Length: 4726  
Score: 2528.00 Matches: 459  
Percent Similarity: 91.03% Conservative: 28  
Best Local Similarity: 85.79% Mismatches: 46  
Query Match: 86.99% Indels: 2  
DB: 14 Gaps: 2

US-09-807-802A-17 (1-534) x AVU48704 (1-4726)

QY 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20  
DB 2815 ATGGCTTCAGCGGTGGCGCACCAATGGCAGACAATAACGAGGGTGCCGATGGAGTGGGT 2874  
QY 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40  
DB 2875 AATTCTCTCAGGAATGGCATTTGCGATTCCCAATGGCTGGCGACAGAGTCATCACCACC 2934  
QY 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
DB 2935 AGCACCAGAACCTGGGCCCTGCCCACTTACAACAACCATCTCTACAAGCAAAATCTCCAGC 2994  
QY 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80  
DB 2995 CAATCA--GGAGCTTCAAACGACAACCACTACTTTGGCTACAGCACCCCTTGGGGGTAT 3051  
QY 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100  
DB 3052 TTTGACTTTAACAGATTCCCACTGCCACTTCTCACACAGTACTGTCGACGAGCTCATTAAC 3111  
QY 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120  
DB 3112 AACAACTGGGATTCCGCCCAAGAACTCAGCTTCAAGCTCTTCAACATCCAAGTTAGA 3171  
QY 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140  
DB 3172 GGGGTACGCAGAACCATGGCACGACGACTATTGCCAATAACCTTACCAGCAGCGTTCAA 3231  
QY 141 valPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160

DB 3232 GTGTTTACGGACTCGGAGTATCAGCTCCCGTACGTCTCGGGTCGGCGCACCAAGGCTGT 3291  
QY 161 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
DB 3292 CTCCCGCGCTTTCCAGCGGACGTCTTTCATGGTCCCTCAGTATGGATACCTCACCCCTGAAC 3351  
QY 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
DB 3352 AACGGAAGTCAAGCGGTGGGACGCTCATCTCTTTACTGCCTGGAGTACTTTCCTTCGAG 3411  
QY 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis 220  
DB 3412 ATGCTAAGGACTGGAAATAAATTCCAATTTCAGCTATACCTTCGAGGATGTACCTTTTCAC 3471  
QY 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
DB 3472 AGCAGTACGCTCACAGCCAGAGTTTGGATCGCTTGATGAATCCTCTTATTGATCAGTAT 3531  
QY 241 LeuTyrTyrLeuAsnArgThrGln--AsnGlnSerGlySerAlaGlnAsnLysAspLeu 259  
DB 3532 CTGTACTACCTGAACAGAACGCAAGGAACAACCTCTGGAACAACCAACCAATCACCGGCTG 3591  
QY 260 LeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGly 279  
DB 3592 CTTTTGTAGCCAGGCTGGGCTCAGTCTATGTCTTTTGAGCCAGAAATTGGCTACCTGGG 3651  
QY 280 ProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPhe 299  
DB 3652 CCCTGCTACCGGCAACAGAGACTTTCAAAGACTGTAAACGACAACAACAACAGTAACCTT 3711  
QY 300 ThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGly 319  
DB 3712 CCTTGGACAGCGGCCAGCAAAATATCATCTCAATGGCGCGACTCGTGGTGAATCCAGGA 3771  
QY 320 ThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMet 339  
DB 3772 CCAGCTATGGCCAGTCCACAAGGACGATGAAGAAAAATTTTCCCTATGCACGGCAATCTA 3831  
QY 340 IlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThr 359  
DB 3832 ATATTGGCAAAGAGGGACAAACGCAAGTAACGCAGAAATTAGATAAATGTAATGATTACG 3891  
QY 360 AspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAla 379  
DB 3892 GATGAAGAAGAGATTTCGTACCACCAATCTCTGTGGCAACAGACGAGTATGGAACCTGTGGCA 3951  
QY 380 ValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAla 399  
DB 3952 AATAACTTCAGAGCTCAATACAGCTCCCACGACTGGAACTGTCAATCATCAGGGGGCC 4011  
QY 400 LeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLys 419  
DB 4012 TTACCTGGCATGGTGTGGCAAGATCGTGACGTGTACCTTCAAGGACCTATCTGGGCAAG 4071  
QY 420 IleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLys 439  
DB 4072 ATTCCTCACACGGATGGACACTTTCATCTCTCTCTCTGATGGGAGGCTTTGGACTGAAA 4131  
QY 440 AsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGlu 459  
DB 4132 CATCCGCTCCTCAAATCATGATCAAAAATACTCCGGTACCGGCAAAATCTCCGACGACT 4191  
QY 460 PheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerVal 479  
DB 4192 TTCAGCCCGGCCAAGTTTTCCTTTCATTTATCACTCAGTACTCCACTGGACAGGTGAGGTG 4251  
QY 480 GluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyr 499  
DB 4252 GAAATTGAGTGGGAGCTACAGAAAGAAAAACAGCAACCGTTGGAATCCAGAGATTCAGTAC 4311  
QY 500 ThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyr 519



Db

4312

ACTTCCAAC

TACAACAG

CTCTGTTA

ATGTGGAC

TTTACTGT

GAGACACT

AAATGGT

TTTAT

4371

Qy

520

ThrGluPro

ArgProIle

GlyThrArg

TyrLeuThr

ArgProLeu

534

Db

4372

AGTGAAC

CTCGCCCT

TATTGGA

ACCGCGT

ATCTCA

CACGAACT

TG

4416

RESULT 5

AR264580

LOCUS

AR264580

DEFINITION

Sequence 5 from patent US 6491907.

ACCESSION

AR264580

VERSION

AR264580.1

GI:29692850

KEYWORDS

Unknown.

SOURCE

Unknown.

ORGANISM

Unknown.

REFERENCE

1 (bases 1 to 8179)

AUTHORS

Rabinowitz,J.E., Samulski,R.J. and Xiao,W.

TITLE

Recombinant parvovirus vectors and method of making

JOURNAL

Patent: US 6491907-A 5 10-DEC-2002;

FEATURES

Location/Qualifiers

source

1..8179

BASE COUNT

2082 a 2078 c 2027 g 1992 t

ORIGIN

/organism="unknown"

Alignment Scores:

Pred. No.:

2.47e-194

Length:

8179

Score:

2518.00

Matches:

458

Percent Similarity:

90.65%

Conservative:

27

Best Local Similarity:

85.61%

Mismatches:

48

Query Match:

86.65%

Indels:

2

DB:

6

Gaps:

2

US-09-807-802A-17 (1-534) x AR264580 (1-8179)

Qy

1

MetAlaSer

GlyGlyGly

AlaProMet

AlaAspAsn

AsnGluGly

AlaAspGly

ValGly

20

Db

2739

ATGGCTAC

AGGCAGT

TGGCGCAC

CAATGGC

AGACAATA

ACGAGG

CGCGCG

GAGTGGGT

2798

Qy

21

AsnAlaSer

GlyAsnTrp

HisCysAsp

SerThrTrp

LeuGlyAsp

ArgValIle

ThrThr

40

Db

2799

AATTCCTC

CGGAAAT

TGGCAT

TGGCAT

TCCAGT

TGGCGC

ACAGAGT

CATCAC

CACC

2858

Qy

41

SerThrArg

ThrTrpAla

LeuProThr

TyrAsn

AsnHisLeu

TyrLysGln

IleSerSer

60

Db

2859

AGCACCGG

AACTGG

CGCCCT

TGCCAC

CTTACA

CAACCA

CCCTCT

TACAAACA

AAATTTCC

AGC

2918

Qy

61

AlaSerThr

GlyAlaSer

AsnAsp

AsnHisTyr

PheGlyTyr

SerThrPro

TrpGlyTyr

80

Db

2919

CAATCA--

GGAGCCT

CGAACG

ACAATCA

CTACTTT

TGGCTA

CAGCAC

CCCCCT

TGGGGTAT

2975

Qy

81

PheAspPhe

AsnArgPhe

HisCysHis

PheSerPro

ArgAspTrp

GlnArgLeu

IleAsn

100

Db

2976

TTTGACTT

CAACAGAT

TCCACTG

CCACTTT

TCCACC

CGTACT

GGCAAG

ACTCAT

CAAC

3035

Qy

101

AsnAsnTrp

GlyPheArg

ProLysArg

LeuAsn

PheLysLeu

PheAsnIle

GlnValLys

120

Db

3036

AACAACT

GGGGATT

CCGACCC

CAAGAG

ACTCAACT

TCAAGCT

CTTTA

ACATTT

CAAGTCA

AA

3095

Qy

121

GluValThr

ThrAsnAsp

GlyValThr

ThrIleAla

AsnAsnLeu

ThrSerThr

ValGln

140

Db

3096

GAGGTAC

GCAGAAT

GACGGT

ACGACG

ACGATT

TGCCAA

TAAACCT

TAC

CAGCAC

CGGTT

CAG

3155

Qy

141

ValPheSer

AspSerGlu

TyrGlnLeu

ProTyrVal

LeuGlySer

AlaHisGln

GlyCys

160

Db

3156

GTGTTACT

GTACTCG

GAGTAC

CGGTACC

AGCTCC

CGTAC

GTGCTCG

GGTCC

GGCGC

ACCAAG

CGTGT

3215

Qy

161

LeuProPro

PheProAla

AspValPhe

MetIlePro

GlnTyrGly

TyrLeuThr

LeuAsn

180

Db

3216

CTCCCGC

CGTTTCC

AGCGG

ACGTTT

CATGGT

CCCTCA

GTATGG

ATACCT

CACCCCT

GAA

3275

Qy

181

AsnGlySer

GlnAlaVal

GlyArgSer

PheTyrCys

LeuGluTyr

PheProSer

Gln

200

Db

3276

AACGGAAG

TCAGCGGT

GGGACG

CTCATCTT

TACTGC

CTGGAG

TACTTCC

CTTCGC

AG

3335

Qy

201

MetLeuArg

ThrGlyAsn

AsnPheThr

PheSerTyr

ThrPheGlu

GluValPro

PheHis

220

Db

3336

ATGCTAAG

GAAGCTG

GAAATAA

CTTCCA

ATTTCAG

CTATAC

CTTCG

AGGATG

TACCTTTT

CAC

3395

Qy

221

SerSerTyr

AlaHisSer

GlnSerLeu

AspArgLeu

MetAsnPro

LeuIleAsp

GlnTyr

240

Db

3396

AGCAGCTA

CGCTCAC

AGCCAG

AGTTTGG

ATCGT

TGATGA

ATCCTCT

TATTAT

GATCAG

TAT

3455

Qy

241

LeuTyrTyr

LeuAsnArg

ThrGln--

AsnGln

SerGlySer

AlaGln

AsnLysAsp

Leu

259

Db

3456

CTGTACTA

CTGAAC

AGAAAC

GGAACAA

CCCTCT

TGGAACA

CAACCA

CAATC

CACGGCTG

3515

Qy

260

LeuPheSer

ArgGlySer

ProAlaGly

MetSerVal

GlnProLys

AsnTrpLeuPro

Gly

279

Db

3516

CTTTTTAG

CCAGGCT

TGGCCTC

AGTCTAT

GTCTTT

GCAGGC

CAGAAAT

TGGCTAC

CTGGG

3575

Qy

280

ProCysTyr

ArgGlnGln

ArgValSer

LysThrLysThr

AspAsn

AsnAsnSer

AsnPhe

299

Db

3576

CCCTGTAC

CCGGCA

ACAGAG

ACTTTCA

AAAGACT

GCTAA

CACACA

CAACAG

TAACTTT

3635

Qy

300

ThrTrpThr

GlyAlaSer

LysTyrAsn

LeuAsnGly

ArgGluSer

IleIleAsn

ProGly

319

Db

3636

CCTTGGAC

AGCGCC

AGCAAT

ATCATCT

CAATG

CGCGG

CACTCG

CTGGTGA

ATCCAGGA

3695

Qy

320

ThrAlaMet

AlaSerHis

LysAspAsp

GluAspLys

PhePhePro

MetSerGly

ValMet

339

Db

3696

CCAGCTAT

TGGCCAGT

CAAGGAC

GATGA

AGAAAA

ATTTTCC

TATGC

ACGCG

CAATCTA

3755

Qy

340

IlePheGly

LysGluSer

AlaGlyAla

SerAsnThr

AlaLeuAsp

AsnValMet

IleThr

359

Db

3756

ATATTTGG

CAAGAG

GGGACA

ACCGCA

AGTAAC

GCAGAA

TTAGATA

ATGTAT

GATTACG

3815

Qy

360

AspGluGlu

GluIleLys

AlaThrAsn

ProValAla

ThrGluArg

PheGlyThr

ValAla

379

Db

3816

GATGAAGA

AGAGATT

CGTACC

CAATCCT

GTGTGC

CAACAG

AGCAGT

ATGGA

AACTGTG

GCAC

3875

Qy

380

ValAsnPhe

GlnSerSer

SerThrAsp

ProAlaThr

GlyAspVal

HisAlaMet

GlyAla

399

Db

3876

AATAACTT

GCAGAG

CTCAA

ATACAG

CTCCC

ACGACT

TGGA

ACTGT

CAATCAT

CATCAGG

GGGCC

3935

Qy

400

LeuProGly

MetValTrp

GlnAspArg

AspValTyr

LeuGlnGly

ProIleTrp

AlaLys

419

Db

3936

TTACCTGG

CATGGT

TGGCAAG

ATCGT

AGCTGT

ACCTT

CAAGG

ACCTAT

CTGGG

CAAA

3995

Qy

420

IleProHis

ThrAspGly

HisPheHis

ProSerPro

LeuMetGly

GlyPheGly

LeuLys

439

Db

3996

ATTCTCAC

ACGATGG

ACACTTT

CATCTT

CTCTCT

CTGTAT

GGGAG

CGCTTT

GGACTG

AA

4055

Qy

440

AsnProPro

ProGlnIle

LeuIleLys

AsnThrPro

ValProAla

AsnProAla

Glu

459

Db

4056

CATCCGCT

CTCAAAT

CATGAT

CAAAA

ATACTCC

GGTAC

CTCGA

ATCCTT

CGACC

CACC

4115

Qy

460

PheSerAla

ThrLysPhe

AlaSerPhe

IleThrGln

TyrSerThr

GlyGlnVal

SerVal

479

Db

4116

TTCAGTGC

GGCAAA

GTGCTT

CTTCAT

CACACAG

TACTCC

ACGGG

ACAGG

TCAGCGT

G

4175

Qy

480

GluIleGlu

TrpGluLeu

GlnLysGlu

AsnSerLys

ArgTrpAsn

ProGluVal

GlnTyr

499

Db

4176

GAGATCGA

TGGGAG

CTGCAG

AGGAA

AAACAG

CAACG

CTGGA

ATCCCC

GAAATTC

AGTAC

4235

Qy

500

ThrSerAsn

TyrAlaLys

SerAlaAsn

ValAspPhe

ThrValAsp

AsnAsnGly

LeuTyr

519

Db

4236

ACTTCCA

ACTACA

CAAGTCT

GTTAAT

CGTGG

ACTTAC

CGTGG

ATACTA

ATGGCGT

GTAT

4295

Qy

520

ThrGluPro

ArgProIle

GlyThrArg

TyrLeuThr

ArgProLeu

534

Db

4296

TCAGAGC

CTCGCC

CAATGG

CACCA

CGATAC

CTGACT

CGTAAT

CTG

4340

RESULT 6

AR222046

LOCUS AR222046 4072 bp DNA linear PAT 26-SEP-2002  
DEFINITION Sequence 4 from patent US 6429001.  
ACCESSION AR222046  
VERSION AR222046.1 GI:23329402  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS 1 (bases 1 to 4072)  
TITLE Hardy,S.F.  
JOURNAL Recombinant AAV packaging systems  
FEATURES Patent: US 6429001-A 4 06-AUG-2002;  
Location/Qualifiers  
1..4072  
/organism="unknown"  
BASE COUNT 1073 a 1025 c 992 g 982 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 9.09e-192 Length: 4072  
Score: 2481.50 Matches: 445  
Percent Similarity: 90.26% Conservative: 37  
Best Local Similarity: 83.33% Mismatches: 51  
Query Match: 85.39% Indels: 1  
DB: 6 Gaps: 1  
US-09-807-802A-17 (1-534) x AR222046 (1-4072)  
Qy 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20  
Db 2090 ATGGCTACAGCGAGTGGCGCACCAATGGCAGACAATAACGAGGCGCCGACGGAGTGGGT 2149  
Qy 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40  
Db 2150 AATTCTCGGAAATTGGCATTGGCATTGCCATGGATGGCGCAGAGTCAATCACCACC 2209  
Qy 41 SerThrArgThrTrpAlaLeuProThrTrpAsnAsnHisLeuTyLysGlnIleSerSer 60  
Db 2210 AGCACCCGAACCTGGGCCCTGCCACCTACAACAACCAACCTCTACAACAATAATTTCCAGC 2269  
Qy 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyLysPheGlyTyLysSerThrProTrpGlyTyL 80  
Db 2270 CAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTGGGGGTAT 2326  
Qy 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100  
Db 2327 TTTGACTTCAACAGATTCCACTGCCACTTTTCAACACGTGACTGGCAAGACTCATCAAC 2386  
Qy 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120  
Db 2387 AACAACTGGGGAATCCGACCCNAGAGACTCAACTTCAAGCTCTTTAAACATTCAAGTCAA 2446  
Qy 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140  
Db 2447 GAGGTCACGCAGAAATGACGGGTACGACGACGATGGCAATTAACCTTACCAGCACGGTTTCA 2506  
Qy 141 ValPheSerAspSerGluTyLysGlnLeuProTyLysValLeuGlySerAlaHisGlnGlyCys 160  
Db 2507 GTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTCTCGGCTCGGCGCATCAAGGATGC 2566  
Qy 161 LeuProProPheProAlaAspValPheMetIleProGlnTyLysGlyTyLysLeuThrLeuAsn 180  
Db 2567 CTCCCGCGGTTCCACGACAGCGTCTTTCATGGTGCCACAGTATGGATACCTCACCCCTGAAC 2626  
Qy 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyLysLeuGluTyLysPheProSerGln 200  
Db 2627 AACGGGAGTCAGGCAGTAGGACGCTCTTCAATTTTACTGCTGGAGTACTTCTCTCTCAG 2686  
Qy 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyLysThrPheGluGluValPropheHis 220  
Db 2687 ATGCTGCGTACCGGAAACAACCTTTACCTTCAGCTACACTTTTGGAGACGTTCTCTTCCAC 2746  
Qy 221 SerSerTyLysAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyL 240

Db 2747 AGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGACCAGTAC 2806  
Qy 241 LeuTyLysLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260  
Db 2807 CTGTATTACTTGAGCAGAAACAAACACTCCAAAGTGGAAACCAACGACGAGTCAAGGCTTCAG 2866  
Qy 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280  
Db 2867 TTTTCTCAGGCCGAGCGAGTGACATTTCGGGACCAGTCTAGGAACCTGGCTTCTCTGGACCC 2926  
Qy 281 CysTyLysArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
Db 2927 TGTACCGCCAGCAGCGAGTATCAAAGACATCTGCGGATAACAACAACAGTGAATACTCG 2986  
Qy 301 TrpThrGlyAlaSerLysTyLysAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320  
Db 2987 TGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGCCCCG 3046  
Qy 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340  
Db 3047 GCCATGGCAAGCCACAGGACGATGAAGAAAAGTTTTCCTCAGAGCGGGGTCTCTCATC 3106  
Qy 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360  
Db 3107 TTTGGGAAGCAAGGCTCAGAGAAAACAATAATGTGGACATTGAAAAGGTCTATGATTACAGAC 3166  
Qy 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380  
Db 3167 GAAGAGGAAATCAGGACAACCAATCCCGTGGCTACGAGCAGTATGGTTCTGTATCTACC 3226  
Qy 381 AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400  
Db 3227 AACCTCCAGAGAGGCAACAGACAAGCAGTACCAGCAGATGTCAACACACAAGCGTCTCT 3286  
Qy 401 ProGlyMetValTrpGlnAspArgAspValTyLysLeuGlnGlyProIleTrpAlaValIle 420  
Db 3287 CCAGGCATGGTCTGGCAGCAGAGATGTGTACTTTCAGGGGCCCATCTGGSCAAAGATT 3346  
Qy 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440  
Db 3347 CCACACACGACGGACATTTTCACCCCTCTCCCTCATGGGTGGATTTCGACCTTAAACAC 3406  
Qy 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460  
Db 3407 CCTCCTCCACAGATTCTCATCAAGAAACACCCCGGTACTCCGCGGACAGGTCAAGCTTC 3466  
Qy 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyLysThrGlyGlnValSerValGlu 480  
Db 3467 AGTCGGGCAAGTTTGTCTTCTTCTCATCACACAGTACTCCACGGGACAGGTCAAGCTGGAG 3526  
Qy 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyLysThr 500  
Db 3527 ATCGAGTGGGAGCTGCAGAAAGAAAACAGCAACCGCTGGAATCCCGAAATTCAGTACACT 3586  
Qy 501 SerAsnTyLysAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyLys 520  
Db 3587 TCCAACACTACAAGTCTGTAAATGTGGACTTTTACTGTGGACACTAATGGCGTGTATTCA 3646  
Qy 521 GluProArgProIleGlyThrArgTyLysLeuThrArgProLeu 534  
Db 3647 GAGCTCGCCCCCATTTGGCACCAGATACCTGACTCGTAACTCG 3688

RESULT 7  
AX205074  
LOCUS AX205074 4072 bp DNA linear PAT 30-AUG-2001  
DEFINITION Sequence 4 from Patent WO0155361.  
ACCESSION AX205074  
VERSION AX205074.1 GI:15394315  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
unclassified.

```
REFERENCE 1
AUTHORS Hardy,S.F.
TITLE Recombinant aav packaging systems
JOURNAL Patent: WO 0155361-A 4 02-AUG-2001;
Chiron Corporation (US)
FEATURES
source 1. .4072
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
/note="recombinant DNA"
BASE COUNT 1073 a 1025 c 992 g 982 t
ORIGIN
Alignment Scores:
Pred. No.: 9.09e-192 Length: 4072
Score: 2481.50 Matches: 445
Percent Similarity: 90.26% Conservative: 37
Best Local Similarity: 83.33% Mismatches: 51
Query Match: 85.39% Indels: 1
DB: 6 Gaps: 1
US-09-807-802A-17 (1-534) x AX205074 (1-4072)
Qy 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20
Db 2090 ATGGCTACAGGCAGTGGGGCACCACCAATGGCAGACAATAACGAGGGCCGACGGAGTGGGT 2149
Qy 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThr 40
Db 2150 AATTCTCGGGAAATTGGCATTTGCCATTCCACATGGATGGGCGACAGAGTCATCACACC 2209
Qy 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60
Db 2210 AGCACCGAACTGGGCCCTGCCACCTACAAACAACCACTTACAAACAATTTCCAGC 2269
Qy 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80
Db 2270 CAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTGGGGGTAT 2326
Qy 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100
Db 2327 TTTGACTTCAACAGATTCCCACTGCCACTTTTCAACCAGTGACTGGCAAGACTCATCAAC 2386
Qy 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120
Db 2387 AACAACTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAAACATTCAAGTCAA 2446
Qy 121 GluValThrThrAsnAspGlyValThrIleAlaAsnAsnLeuThrSerThrValGln 140
Db 2447 GAGGTACCGCAGAAATGACGGTACGACGACGATTGCCAATAACCTTACCAGCACGGTTCCAG 2506
Qy 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160
Db 2507 GTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTCTCGGCTCGGCGCATCAAGGATGC 2566
Qy 161 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180
Db 2567 CTCCCGCGGTTCACAGCAGACGCTCTTCATGGTGCCACAGTATGGATACCTCACCCCTGAAC 2626
Qy 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200
Db 2627 AACGGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCGCTGGAGTACTTTCTCTTCAG 2686
Qy 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis 220
Db 2687 ATGCTGCGTACCGGAAACAACATTTACCTTCAGCTACACTTTTGGAGGACGTTCTCTTTCCAC 2746
Qy 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240
Db 2747 AGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCTCTCATCGACCACTAC 2806
Qy 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260
```

```
Db 2807 CTGTATTACTTGAGCAGAAACAACACTCCAAGTGGAAACCACCACGCGAGTCAAGGCTTCAG 2866
Qy 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280
Db 2867 TTTTCTCAGGCCGAGCGAGTGACATTCCGGACCAGTCTAGGAACCTGGCTTCCTCGGACCC 2926
Qy 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300
Db 2927 TGTTACCGCCAGCAGCGAGTATCAAAGACATCTGCGGATAACAACAACAGTGAATACTCG 2986
Qy 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320
Db 2987 TGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGCCCCG 3046
Qy 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340
Db 3047 GCCATGGCAAGCCACAGGACGATGAAGAAAGTTTTCCTCAGAGCGGGTCTTCATC 3106
Qy 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360
Db 3107 TTTGGGAAGCAAGGCTCAGAGAAACAATAATGTGGACATTGAAAAGGTCAATGATTACAGAC 3166
Qy 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380
Db 3167 GAAAGGAAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTATGGTTCTGTATCTACC 3226
Qy 381 AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400
Db 3227 AACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAACACACACAAGGCGTCTT 3286
Qy 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420
Db 3287 CCAGGCATGGTCTGGCAGGACAGAGATGTGTACTTTCAGGGGCCCATCTGGGCAAGATT 3346
Qy 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440
Db 3347 CCACACACGGACGGACATTTCACCCCTCTCCCTCATGGGTGGATTTCGGACTTTAAACAC 3406
Qy 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460
Db 3407 CTTCTCCACAGATTCTCATCAAGAACACCCCGTACCTCGGAATCCTTCGACCACCTTC 3466
Qy 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480
Db 3467 AGTCCGGCAAAGTTTGCTTCTTCATCACACAGTACTCCACGGGACAGGTCAGCGTGGAG 3526
Qy 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 500
Db 3527 ATCAGTGGGAGCTGCAGAAGGAAACAGACAAACGCTGGAATCCCGAAATTCAGTACACT 3586
Qy 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520
Db 3587 TCCAACTACAACAAGTCTGTTAATGTGGACTTTACTGTGGACACTAATGGCGGTATTCA 3646
Qy 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534
Db 3647 GAGCCTCGCCCCCATTTGGCACCCAGATACCTGACTCGTAATCTG 3688
RESULT 8
AX282480
LOCUS AX282480 4679 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 1 from Patent WO0168888.
ACCESSION AX282480
VERSION AX282480.1 GI:16609611
KEYWORDS Adeno-associated virus 2
SOURCE Adeno-associated virus 2
ORGANISM Adeno-associated virus 2
REFERENCE 1
AUTHORS Xiao,W. and During,M.J.
TITLE Production of chimeric capsid vectors
JOURNAL Patent: WO 0168888-A 1 20-SEP-2001;
```



|   |                            |   |      |
|---|----------------------------|---|------|
| Neurologix, Inc. (US)                         |                            | Location/Qualifiers   |      |
| source  | 1..4679                    |   |      |
| /organism="Adeno-associated virus 2"          |                            |   |      |
| /mol_type="genomic DNA"                       |                            |   |      |
| /db_xref="taxon:10804"                        |                            |   |      |
| BASE COUNT                                    | 1198 a 1262 c 1255 g 964 t |   |      |
| ORIGIN  |                            |   |      |
| Alignment Scores:                             |                            |   |      |
| Pred. No.:                                    | 1.09e-191                  | Length:   | 4679 |
| Score:  | 2481.50                    | Matches:  | 445  |
| Percent Similarity:                           | 90.26%                     | Conservative:   | 37   |
| Best Local Similarity:                        | 83.33%                     | Mismatches:   | 51   |
| Query Match:                                  | 85.39%                     | Indels:   | 1    |
| DB:   | 6                          | Gaps:   | 1    |
| US-09-807-802A-17 (1-534) x AX282480 (1-4679) |                            |   |      |
| QY  | 1                          | MetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly    | 20   |
| Db  | 2809                       | ATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGCGCCGACGGAGTGGGT     | 2868 |
| QY  | 21                         | AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr    | 40   |
| Db  | 2869                       | AATTCTCGGGAATTGGCATTGGCATTTCCATGATGGCGCACAGATCATCACCACC         | 2928 |
| QY  | 41                         | SerThrArgThrTrpAlaLeuProThrTyraAsnAsnHisLeuTyryLysGlnIleSerSer  | 60   |
| Db  | 2929                       | AGCACCGGAACCTGGGCCCTGCCACCTACAACAACCCACCTCTACAAACAATTTCCAGC     | 2988 |
| QY  | 61                         | AlaSerThrGlyAlaSerAsnAspAsnHisTyryPheGlyTyrySerThrProTrpGlyTyry | 80   |
| Db  | 2989                       | CAATCA---GGAGCCTCGAACGACAATCACTATTGGCTACAGCACCCCTTGGGGGTAT      | 3045 |
| QY  | 81                         | PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn    | 100  |
| Db  | 3046                       | TTTGACTTCAACAGATTCCACTGCCACTTTTCCACCACGTGACTGGCAAGACTCATCAAC    | 3105 |
| QY  | 101                        | AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys    | 120  |
| Db  | 3106                       | AACAACCTGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAATCAAGTCAAA      | 3165 |
| QY  | 121                        | GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln    | 140  |
| Db  | 3166                       | GAGGTACACGCAAGTACGGTACGACGACGATTGCCAATAACCTTACCAGCACGGTTTACG    | 3225 |
| QY  | 141                        | ValPheSerAspSerGluTyryGlnLeuProTyryValLeuGlySerAlaHisGlnGlyCys  | 160  |
| Db  | 3226                       | GTGTTTACTGACTCGGAGTACCAGTCCCGTACGTCTCTCGGCTCGGCGCATCAAGGATGC    | 3285 |
| QY  | 161                        | LeuProProPheProAlaAspValPheMetIleProGlnTyryGlyTyryLeuThrLeuAsn  | 180  |
| Db  | 3286                       | CTCCCGCGTTCCAGCAGACGTCTTTCATGGTGCCACAGTATGGATACCTCACCCCTGAAC    | 3345 |
| QY  | 181                        | AsnGlySerGlnAlaValGlyArgSerSerPheTyryCysLeuGluTyryPheProSerGln  | 200  |
| Db  | 3346                       | AACGGGAGTCAGGCAGTAGGACGCTCTTCAATTTACTGCCTGGAGTACTTTCTTCTCAG     | 3405 |
| QY  | 201                        | MetLeuArgThrGlyAsnAsnPheThrPheSerTyryThrPheGluGluValPropheHis   | 220  |
| Db  | 3406                       | ATGCTGCGTACCGGAACAACACTTTACCTTCAGCTACACTTTTGGAGGACGTTCTTTCCAC   | 3465 |
| QY  | 221                        | SerSerTyryAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyry  | 240  |
| Db  | 3466                       | AGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGACCAGTAC    | 3525 |
| QY  | 241                        | LeuTyryTyryLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu  | 260  |
| Db  | 3526                       | CTGTATTACTTGAGCAGACAACAACACTCCAAGTGGGAACCAACCGCAGTCAGGCTTCAG    | 3585 |
| QY  | 261                        | PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro    | 280  |

|            |   |   |                                    |
|------------|---|---|------------------------------------|
| Db         | 3586  | TTTTTCTCAGCCGCGAGCGAGTGACATTCCGGACCAGTCTAGGAACCTGGCTTCTCTGGACCC | 3645                               |
| QY         | 281   | CysTyryArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr   | 300                                |
| Db         | 3646  | TGTTACCGCCAGCAGCGAGTATCAAAGACATCTCGCGATAACAACAACAGTGAATACTCG    | 3705                               |
| QY         | 301   | TrpThrGlyAlaSerLysTyryAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr   | 320                                |
| Db         | 3706  | TGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGCCCCG   | 3765                               |
| QY         | 321   | AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle    | 340                                |
| Db         | 3766  | GCCATGGCAAGCCACAGGACGATGAAGAAAAGTTTTCCTCAGAGCGGGGTTCTCATC       | 3825                               |
| QY         | 341   | PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp    | 360                                |
| Db         | 3826  | TTTGGGAAGCAAGGCTCAGAGAAAACAATGTGGACATTGAAAAGGTCTATGATTACAGAC    | 3885                               |
| QY         | 361   | GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal    | 380                                |
| Db         | 3886  | GAAGAGGAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTATGGTTCTGTATCTACC     | 3945                               |
| QY         | 381   | AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu    | 400                                |
| Db         | 3946  | AACCTCCAGAGAGGCAACAGACAAGCAGTACCAGGATGTCAACACACAACAGGCGTTCTT    | 4005                               |
| QY         | 401   | ProGlyMetValTrpGlnAspArgAspValTyryLeuGlnGlyProIleTrpAlaIle      | 420                                |
| Db         | 4006  | CCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGCAAGATT     | 4065                               |
| QY         | 421   | ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn    | 440                                |
| Db         | 4066  | CCACACACGACGGACATTTTCAACCCCTCTCCCTCATGGGTGGATTCCGACTTAAACAC     | 4125                               |
| QY         | 441   | ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe    | 460                                |
| Db         | 4126  | CCTCTCCACAGATTCTCATCAAGAAACACCCCGGTACTCTGCGAATCTCTCGACCACCTTC   | 4185                               |
| QY         | 461   | SerAlaThrLysPheAlaSerPheIleThrGlnTyrySerThrGlyGlnValSerValGlu   | 480                                |
| Db         | 4186  | AGTGGGCAAAAGTTTGCTTCTCTTCTCATCACACAGTACTCCACGGGACAGGTCAGCGTGGAG | 4245                               |
| QY         | 481   | IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyryThr   | 500                                |
| Db         | 4246  | ATCGAGTGGGAGCTGCAGAGGAAACAGCAACGCTGGATCCCGAAATTCAGTACACT        | 4305                               |
| QY         | 501   | SerAsnTyryAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyryThr  | 520                                |
| Db         | 4306  | TCCAACTACAACAGTCTGTTAATGTGGACTTTTACTGTGGACACTAATGGCGTGTATTCA    | 4365                               |
| QY         | 521   | GluProArgProIleGlyThrArgTyryLeuThrArgProLeu                     | 534                                |
| Db         | 4366  | GAGCCTCGCCCCATTGGCACCAGATACCTGACTCGTAATCTG                      | 4407                               |
| RESULT 9   |   |   |                                    |
| AF043303   | LOCUS   | AF043303  | 4679 bp DNA linear VRL 24-FEB-1998 |
| DEFINITION | Adeno-associated virus 2, complete genome.                        |   |                                    |
| ACCESSION  | AF043303  |   |                                    |
| VERSION    | AF043303.1  | GI:2906016  |                                    |
| KEYWORDS   |   |   |                                    |
| SOURCE     | Adeno-associated virus 2  |   |                                    |
| ORGANISM   | Adeno-associated virus 2  |   |                                    |
| REFERENCE  | Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus. |   |                                    |
| AUTHORS    | 1 (bases 1 to 4679)   |   |                                    |
| TITLE      | Ruffing,M., Heid,H. and Kleinschmidt,J.A.                         |   |                                    |
|            | Mutations in the carboxy terminus of adeno-associated virus 2     |   |                                    |
|            | capsid proteins affect viral infectivity: lack of an RGD          |   |                                    |
|            | integrin-binding motif  |   |                                    |
| JOURNAL    | J. Gen. Virol. 75 (Pt 12), 3385-3392 (1994)                       |   |                                    |
| MEDLINE    | 95088582  |   |                                    |
| PUBMED     | 7996133   |   |                                    |
| REFERENCE  | 2 (bases 1 to 4679)   |   |                                    |

AUTHORS Berns,K.I., Bohenzky,R.A., Cassinotti,P., Colvin,D., Donahue,B.A., Dull,T., Horer,M., Kleinschmidt,J.A., Ruffing,M., Snyder,R.O., Tratschin,J.-D. and Weitz,M.

TITLE Direct Submission

JOURNAL Submitted (15-JAN-1998) Cell Genesys Inc., 342 Lakeside Dr., Foster City, CA 94404, USA

FEATURES

source

1. .4679

/organism="Adeno-associated virus 2"

/mol\_type="genomic DNA"

/db\_xref="taxon:10804"

/note="changes relative to the original sequence, GenBank Accession Number J01901, have been detected and verified by several different laboratories"

repeat\_region

1. .145

/note="inverted terminal repeat"

/rpt\_type=inverted

misc\_feature

42. .83

/note="flip oriented DNA"

precursor\_RNA

287. .4451

join(321. .1906,2228. .2252)

/codon\_start=1

/product="Rep 68 protein"

/protein\_id="AAC03774.1"

/db\_xref="GI:2906017"

/translation="MPGFYEIVIKVPSDDLDEHLPGISDSFVNWVAEKEWELPPDSMD LNLIEQAPLTVAEKLQRLDFTWRRVSKAPEALFFVQFEKGESYFHMHLVETTGVS MVLGRFLSQIREKLIQRIYRGIEPTLPNWFAVTKTRNGAGGKNKVDECIYPNLLPK TQPELQAWTNMEQYLSACLNLTERKRLVAQHLTHVSQTQEONKENQNPSDAPVIRS KTSARYMELVGLVDKGITSEKQWIQEDQASYISFNAAASNSRSQIKAALDNAGKIMSL TKTAPDYLVGQQPVEDISSNRIYKILELNGYDPQYAAASVFLGWATKKFGKRTIWLFG PATTGKTNIAEIAHTVPFYGCNVNTNENFPFNDQVDKMWIWEEGKMTAKVVESAKA ILGSKVRVDQKCKSSAQIDPTPVIVTSNTNMCVIDGNSTTTFEHQQPLQDRMFKFEL TRRLDHDGFKVTKQEVKDFFRWAKOHVVEHEFVYVKGGAKKRPAPSDADISEPKRV RESVAQPSTSDAEASINYADRLARGHSL"

321. .2186

/codon\_start=1

/product="Rep 78 protein"

/protein\_id="AAC03775.1"

/db\_xref="GI:2906018"

/translation="MPGFYEIVIKVPSDDLDEHLPGISDSFVNWVAEKEWELPPDSMD LNLIEQAPLTVAEKLQRLDFTWRRVSKAPEALFFVQFEKGESYFHMHLVETTGVS MVLGRFLSQIREKLIQRIYRGIEPTLPNWFAVTKTRNGAGGKNKVDECIYPNLLPK TQPELQAWTNMEQYLSACLNLTERKRLVAQHLTHVSQTQEONKENQNPSDAPVIRS KTSARYMELVGLVDKGITSEKQWIQEDQASYISFNAAASNSRSQIKAALDNAGKIMSL TKTAPDYLVGQQPVEDISSNRIYKILELNGYDPQYAAASVFLGWATKKFGKRTIWLFG PATTGKTNIAEIAHTVPFYGCNVNTNENFPFNDQVDKMWIWEEGKMTAKVVESAKA ILGSKVRVDQKCKSSAQIDPTPVIVTSNTNMCVIDGNSTTTFEHQQPLQDRMFKFEL TRRLDHDGFKVTKQEVKDFFRWAKOHVVEHEFVYVKGGAKKRPAPSDADISEPKRV RESVAQPSTSDAEASINYADRYQNKCSRHVGMNMLFPPCRQERMNQNSNICFTHGQK DCLECFPVSESQPVSVVKAYQKLCYIHIMGKVPDACTACDLVNVLDLDDCIFEQ"

370

/note="compared to sequence of GenBank Accession Number J01901"

/replace="g"

878. .4451

join(993. .1906,2228. .2252)

/codon\_start=1

/product="Rep 40 protein"

/protein\_id="AAC03776.1"

/db\_xref="GI:2906019"

/translation="MELVGWLVDKGITSEKQNIQEDQASYISFNAAASNSRSQIKAAL DNAGKIMSLTKTAPDYLVGQQPVEDISSNRIYKILELNGYDPQYAAASVFLGWATKKFGK RNTIWLFGPATTGKTNIAEIAHTVPFYGCNVNTNENFPFNDQVDKMWIWEEGKMTA KVVESAKAILGGSKVRVDQKCKSSAQIDPTPVIVTSNTNMCVIDGNSTTTFEHQQPLQ DRMFKPELTRRLDHDGFKVTKQEVKDFFRWAKOHVVEHEFVYVKGGAKKRPAPSDA DISEPKRVRESVAQPSTSDAEASINYADRLARGHSL"

993. .2186

/codon\_start=1

/product="Rep 52 protein"

/protein\_id="AAC03777.1"

/db\_xref="GI:2906020"

/translation="MELVGWLVDKGITSEKQNIQEDQASYISFNAAASNSRSQIKAALD

NAGKIMSLTKTAPDYLVGQQPVEDISSNRIYKILELNGYDPQYAAASVFLGWATKKFGK RNTIWLFGPATTGKTNIAEIAHTVPFYGCNVNTNENFPFNDQVDKMWIWEEGKMTA KVVESAKAILGGSKVRVDQKCKSSAQIDPTPVIVTSNTNMCVIDGNSTTTFEHQQPLQ DRMFKPELTRRLDHDGFKVTKQEVKDFFRWAKOHVVEHEFVYVKGGAKKRPAPSDA DISEPKRVRESVAQPSTSDAEASINYADRYQNKCSRHVGMNMLFPPCRQERMNQNSN ICFTHGQKDCLECFPVSESQPVSVVKAYQKLCYIHIMGKVPDACTACDLVNVLDLDD CIFEQ"

precursor\_RNA

1853. .4451

intron

1907. .2227

2203. .4410

/codon\_start=1

/product="major coat protein VP1"

/protein\_id="AAC03780.1"

/db\_xref="GI:2906023"

/translation="MAADGYLPDWLEDTLSEGIRQWVKLKPGRPPPKPAERHKDDSRG LVLPGYKYLGPFGNLDKGEPVNEADAALEHDKAYDRQLDSGDNPYLKYNHADAEFQE RLKEDTSFGNLRGRAVFQAKRVLLEPLGVEEPVKTAPGKKRPVEHSPVEPDSSSGTG KAGQOPARKRLNFGQTGDADSVDPQPLGQPPAAPSGLTNTMATGSGAPMADNNEGA DVGNSSGNWHCDSTWMGDRVITTSRTWALPTYNNHLYKQISSQSGASNDNHYFGYS TPWGYDFNRPFCHFSPRDQRLINNNWGFPRKRLNFKLFNIOQKEVTQNDGTTTIAN NLTSTVOVFTDSEYQLPYVLGSAHQCLPPFPVAHSPQYQVGYLTLLNNGSAQVGRSSP YCLEYFPSQMLRTGNFTFSYTFEDVPFHSSYAHSQVRLDRLMNPLIDQLYLSRTNT PSGTTTQSRLLQFSQAGSADIRDOSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKY HLNGRDSLVPNGPAMASHKDDEKFFPSGVLIFGKQSEKTNVDIEKVMITDEEIR TTNPVATEQYGSVSTNLQRNROAATADVNTQGVLPGMVWQDRDVLQGPIMWAKIPHT DGHFHPSPLMGGFGLKHPPPOILLIKNTPVPANPSTTFSAAKPAFSTITQYSTQGVSV EIWELQENSKRWNPETIQYTSNYSNVVDFTVDTNGVYSEPRPIGTRYLTRNL"

2429

/note="compared to sequence of GenBank Accession Number J01901"

/replace="ta"

2614. .4410

/codon\_start=1

/transl\_except=(pos:2614. .2616,aa:Met)

/product="major coat protein VP2"

/protein\_id="AAC03778.1"

/db\_xref="GI:2906021"

/translation="MAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADS VPDQPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRV IITSTRWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYDFNRFHCHFSPRDWQ RLINNNWGFPRKRLNFKLFNIOQKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLG SAHQCLPPFPADVFMPVQYGYLTLLNNGSAVGRSSFYCLEYFPSQMLRTGNNTFSY TFEDEVPHSSYAHSQVSLDRLMNPLIDQLYLSRTNTPSGTTTQSRLLQFSQAGASHDD DQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDD EEKFFPSGVLIFGKQSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQRGN RQAATADVNTQGVLPGMVWQDRDVLQGPIMWAKIPHTDGHFHPSPLMGGFGLKHPPPO ILLIKNTPVPANPSTTFSAAKPAFSTITQYSTQGVSVIEWELQENSKRWNPETIQYTSN YNKSNVVDFTVDTNGVYSEPRPIGTRYLTRNL"

2809. .4410

/codon\_start=1

/product="major coat protein VP3"

/protein\_id="AAC03779.1"

/db\_xref="GI:2906022"

/translation="MATSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSRT WALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYDFNRFHCHFSPRDWQRLINNNW GFRPKRLNFKLFNIOQKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQCL PFPADVFMPVQYGYLTLLNNGSAVGRSSFYCLEYFPSQMLRTGNNTFSYTFEDVFP HSSYAHSQSLDRLMNPLIDQLYLSRTNTPSGTTTQSRLLQFSQAGASDIRDQSRNWL PGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDEKFFPQ SGVLIFGKQSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQRNROAATAD VNTQGVLPGMVWQDRDVLQGPIMWAKIPHTDGHFHPSPLMGGFGLKHPPPOILLIKNTP VPANPSTTFSAAKPAFSTITQYSTQGVSVIEWELQENSKRWNPETIQYTSNYSNVVD FTVDTNGVYSEPRPIGTRYLTRNL"

2877

/note="compared to sequence of GenBank Accession Number J01901"

/replace="c"

3759. .3765

/note="compared to sequence of GenBank Accession Number J01901"

/replace="g"

3859

/note="compared to sequence of GenBank Accession Number

```

J01901"
/replace="a"
3898
/notes="compared to sequence of GenBank Accession Number
J01901"
/replace="a"
3900..3902
/notes="compared to sequence of GenBank Accession Number
J01901"
/replace="gaac"
4232..4233
/notes="compared to sequence of GenBank Accession Number
J01901"
/replace="acg"
4329..4330
/notes="compared to sequence of GenBank Accession Number
J01901"
/replace="tcg"
4336
/notes="compared to sequence of GenBank Accession Number
J01901"
/replace="c"
4341
/notes="compared to sequence of GenBank Accession Number
J01901"
/replace="t"
4535..4679
/notes="inverted terminal repeat"
/rpt type=inverted
4597..4638
/notes="flop oriented DNA"
BASE COUNT 1198 a 1262 c 1255 g 964 t
ORIGIN

Alignment Scores:
Pred. No.: 1.09e-191 Length: 4679
Score: 2481.50 Matches: 445
Percent Similarity: 90.26% Conservative: 37
Best Local Similarity: 83.33% Mismatches: 51
Query Match: 85.39% Indels: 1
DB: 14 Gaps: 1

US-09-807-802A-17 (1-534) x AF043303 (1-4679)
QY 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20
Db 2809 ATGGCTACAGGCAGTGGCGCACCAATGGCAGACATAACGAGGGCGCCGAGTGGGT 2868
QY 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40
Db 2869 AATTCTCGGGAAATTGGCATTTGGCATTTCCGATTCCACATGGATGGCGACAGATCATCCACC 2928
QY 41 SerThrArgThrTrpAlaLeuProThrTrpAsnAsnHisLeuTyrLysGlnIleSerSer 60
Db 2929 AGCACCGGAACCTGGGCGCTGCCACCTACAAACACCACTCTACAAACAAATTTCCAGC 2988
QY 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80
Db 2989 CAATCA--GGAGCCTCGAACGACAAATCACTACTTTTGGCTACAGCACCCCTTGGGGGTAT 3045
QY 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100
Db 3046 TTTGACTTCAACAGATTCCACTGCCACTTTTTCACCACGTGACTGGCAAGACTCATCAAC 3105
QY 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120
Db 3106 AACAACTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAACATTCAAGTCAA 3165
QY 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140
```

```

Db 3166 GAGGTACGCAGAAATGACGGTACGACGACGAGATTGCCAATAACCTTACCAGCACGGTTTCAG 3225
QY 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160
Db 3226 GTGTTTACTGACTCGGAGTACGACGCTCCCGTACGTCTCCGCTCGGCGCATCAAGGATGC 3285
QY 161 LeuProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180
Db 3286 CTCCCGCGCTTCCAGCAGACGCTCTTCATTTTACTGCCTGGAGTACCTTCCCTTCTCAG 3345
QY 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200
Db 3346 AACGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCCTGGAGTACCTTCCCTTCTCAG 3405
QY 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPropheHis 220
Db 3406 ATGCTGCGTACCGGAACAACCTTACCTTCAGCTACACITTTTGAGGACGTTCTTCTCCAC 3465
QY 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240
Db 3466 AGCAGTACGCTCACAGCCAGAGTCTGGACCGTCTCATGATGATCTCTCATCGACCATAC 3525
QY 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260
Db 3526 CTGTATTACTTGAGCAGACAACACTCCAAGTGGAAACCACCGCAGTCAAGGCTTCAG 3585
QY 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280
Db 3586 TTTTCTCAGGCGGAGCGAGTGACATTCGGGACCAGTCTAGGAACCTGGCTTCTGGACCC 3645
QY 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300
Db 3646 TGTTACCGCCAGCAGCGAGTATCAAAGACATCTGCGGATAACAACAACAGTGAATACTCG 3705
QY 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320
Db 3706 TGGACTGGAGTACCAAGTACCCTCAATGGCAGAGACTCTCTGGTGAATCCGGGCGCCG 3765
QY 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340
Db 3766 GCCATGGCAAGCCACAAGGACGATGAAGAAAGTTTTTCTCAGAGCGGGTTCTCATC 3825
QY 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360
Db 3826 TTTGGGAAGCAAGGCTCAGAGAAACAATAATGTGGACATTTGAAAAGGTATGATTACAGAC 3885
QY 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380
Db 3886 GAAGAGGAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTATGGTTCTGTATCTACC 3945
QY 381 AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400
Db 3946 AACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAACACACACAAGGCTTCT 4005
QY 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420
Db 4006 CCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGCAAGATT 4065
QY 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440
Db 4066 CCACACACGACGACGACATTTTTCACCCCTCTCCCTCATGGGTGGATTTCGACTTAAACAC 4125
QY 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460
Db 4126 CCTCCTCCACAGATTCTCATCAAGAAACACCCCGGTACCTTCGGAATCCTTCGACCACCTTC 4185
QY 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480
Db 4186 AGTGGCGCAAGATTGCTTCTTCTTCTCATCACACAGTACTCCACGGGACAGGTCAGCGTGGAG 4245
QY 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 500
```



Dbb 4246 ATCGAGTGGGAGCTGCAGAGGAAAAACAGCAACCGCTGGAATCCGAAATTCAGTACACT 4305

Qy 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520  
|||||  
|||||

Dbb 4306 TCCAACTACAACAAGTCTGTTAATGTGGACTTTTACTGTGGACACTAATATGGCGTGTATTCA 4365

Qy 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
|||||  
|||||

Dbb 4366 GAGCCTCGCCCATTTGGCACCAGACATACCTGACTCGTAATCTG 4407

RESULT 10

AF369963

LOCUS AF369963 7327 bp DNA circular SYN 22-JUL-2001

DEFINITION Cloning vector pAAV-RC, complete sequence.

ACCESSION AF369963

VERSION AF369963.1 GI:14994089

KEYWORDS

SOURCE Cloning vector pAAV-RC

ORGANISM Cloning vector pAAV-RC

REFERENCE 1 (bases 1 to 7327)

AUTHORS Grafsky,A.J. III.

TITLE pAAV-RC for AAV Helper-Free System

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 7327)

AUTHORS Grafsky,A.J. III.

TITLE Direct Submission

JOURNAL Submitted (16-APR-2001) Technical Services, Stratagene, 11011 N. Torrey Pines Rd., La Jolla, CA 92037, USA

FEATURES

Location/Qualifiers

1..7327

/organism="Cloning vector pAAV-RC"

/mol\_type="genomic DNA"

/db\_xref="taxon:161364"

/lab\_host="Escherichia coli"

/note="used for recombinant DNA processing"

131..1996

/note="AAV-2 rep"

/codon\_start=1

/transl\_table=11

/product="adeno-associated virus 2 rep protein"

/protein\_id="AAK76418.1"

/db\_xref="GI:14994090"

CDS

translation="MPGFYEIVIKVPSDLDEHLPGISDSFVNWVAEKWELPPDSMD  
LNLI EQAPLTVAEKLODFLTWRVRVSKAPEALFVQPEKGESYFHMVLTETGVKS  
MVLGRFLSQIREKLIQRIYRGIEPTLPNFWAVTKTRNGAGGKNKVDECIYINLLPK  
TQPELQAWTNMEQVLSACLNLTERKRLVAQHLTHVSQTQEONKENQNSDAPVRS  
KTSARYMELVGLVDKGITSEKQWIQEDQASYISFNAASNSRIKALDNAGKIMSL  
TKTAPDYLVGQQVEDISSNRIYKILELNGYDPOYAAVSFLGWATKKFGKRNITWLFQ  
PATGTKNIAEIAHTVPFYGCNVNTNENFPENDCVDMKVIWEEGRMTAKVVESAKA  
ILGSKVRVDQCKSSAQIDPTPVI VTSNTMCAVIDGNSTTFEHQQPLQDRMFKFEL  
TRRLDHDVGKVTQKEVKDFRWAQKHVVEHEFYPVVKGGAKRPAAPSADISEPKRV  
RESVAQPTSDAEASINYADRYQNKCSRHVGNLMLFPCRQERNQNSNICFTHGQK  
DCLECFPVSESPVSVVKAYQKLCYIIHHIMGRKVPDACTACDLVNVLDLDDCI FEQ"

2013..4220

CDS

/note="cap vp1"

/codon\_start=1

/transl\_table=11

/product="capsid protein VP1"

/protein\_id="AAK76419.1"

/db\_xref="GI:14994091"

translation="MAADGYLPDWLEDTLSEGRQWVKLKPGRPPPPKPAERHKDDSRG  
LVLPQYKYLGPFGNLDKEPVPNEADAALEHDKAYDRQLDSDGNPYLKYKNHADAEFQE  
RLKEDTSFGNLRGAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVHSPEVPDSSSGTG  
KAGQAPARKLRFQGTGDADSVDPDQPLGQPPAAPSGLGTNTMATGSGAPMADNNEGA  
DGVGNSGNHWCSDTWMGDRVITSTRTWALPTYNHLYKQISSQSGASNDNHYFGYS  
TPWGYDFNRFHCHFSPRDWQRLLINNWGFRPKRLNFKLNIQVKEVTQNDGTTTIAN  
NLSTVQVFTDSEYQLPYLGSAHQGLPPFPADFMVFPQYGLTLNNGSQAAGRSSF  
YCLEYFPQMLRTGNFTFSYTFEDVPFHSSYAHQSGLDRLMNPLIDQLYLYLSRTNT  
PSGTTQSLRQFSQAGASDIRQSNWLPGPCYRQSVSKTSADNNNSYSWTGATKY  
HLNGRDSLVPNGPAMASHKDDDEKFFPQSGVLI FGQKSEKTNVDIEKVNITDEEIR  
TTPVPATEQYGSVSTNLQRGNRQAATADVNTQGLVPGMWQDRDVYLOQPIWAKI PHT  
DGHFHPSPLMGGFGLKHPPPPQILIKNTVPVNPSTTFSAAKFAFITQYSTGTQVSV EI

rep\_origin

4705..5161

/note="f1 origin"

CDS

5292..6152

/note="confers ampicillin resistance"

/codon\_start=1

/transl\_table=11

/product="beta-lactamase"

/protein\_id="AAK76420.1"

/db\_xref="GI:14994092"

translation="MSIQHFRVALIPFFFAFCLPVFAHPETLVKVKDAEDQLGARVGY  
IELDLSGKILESFRPEERFPMSTFKVLLCGAVLSRIDAGQEQLGRRIRHYSQNDLVE  
YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTIGGPKELTAFLHNMGDHVTSL  
DRWEPELNEAIPNDERDITMPVAMATTLRKLTGELLTLASRQQLIDWMEADKAVAGPL  
LRSALPAGWFIADKSGAGERSGRIIAALGPDGKPSRIVVIYTTGSAQTMDERNRQIA  
EIGASLIKHW"

6161..7019

/note="Cole1 origin"

BASE COUNT 1903 a 1887 c 1912 g 1625 t

ORIGIN

Alignment Scores:

Pred. No.: 1.98e-191 Length: 7327

Score: 2481.50 Matches: 445

Percent Similarity: 90.26% Conservative: 37

Best Local Similarity: 83.33% Mismatches: 51

Query Match: 85.39% Indels: 1

DB: 12 Gaps: 1

US-09-807-802A-17 (1-534) x AF369963 (1-7327)

Qy 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20  
|||||

Dbb 2619 ATGGCTACAGCGAGTGGCGCACCAATGGCAGACAATAACGAGGGCGCCGAGAGTGGGT 2678

Qy 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40  
|||||

Dbb 2679 AATTCTCGGGAATTGGCATTGCGATTCCACATGGATGGCGACAGAGTCATCACCACC 2738

Qy 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
|||||

Dbb 2739 AGCACCCGAACCTGGGCCCTGCCACCTACACAACACCCTCTACAAACAATTTCCAGC 2798

Qy 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80  
|||||

Dbb 2799 CAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTGGGGGTAT 2855

Qy 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100  
|||||

Dbb 2856 TTTGACTTCAACAGATTCCACTGCCACTTTTACCACCGTGA CTGGCAAGACTCATCAAC 2915

Qy 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120  
|||||

Dbb 2916 AACAACTGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAACAATTCAGTCAAA 2975

Qy 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140  
|||||

Dbb 2976 GAGGTCACGCAGAAATGACGGTACGACGACGATTGCCAATAACCTTACCAGCAGGTTTCAG 3035

Qy 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160  
|||||

Dbb 3036 GTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTCTCGGCTCGGCGCATCAAGGATGC 3095

Qy 161 LeuProPropheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
|||||

Dbb 3096 CTCCCGCCGTTTCCAGCAGACGCTCTTCATGGTGCACAGTATGGATACCTCACCCCTGAAC 3155

Qy 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
|||||

Dbb 3156 AACGGGAGTCAGGAGTAGGACGCTCTTCAATTTTACTGCCTGGAGTACTTTCTCTCTCAG 3215

Qy 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis 220  
|||||

Db 3216 ATGCTGCGTACCGGAAACAACACTTTACCTTCAGCTACACTTTTGAGGACGTTCTCTTCCAC 3275

Qy 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240

Db 3276 AGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGACCATAC 3335

Qy 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260

Db 3336 CTGTATTACTTGAGCAGAAACAACACTCCAAGTGGAAACCAACCAACGACGCTCAAGGCTTCAG 3395

Qy 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280

Db 3396 TTTTCTCAGGCGGAGCGAGTGACATTCGGGACCATCTAGGAACCTGGCTTCTCTGACCC 3455

Qy 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300

Db 3456 TGTTACGCGCAGCAGCGAGTATCAAAAGACATCTGCGGATAACAACAACAGTGAATACTCG 3515

Qy 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320

Db 3516 TGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGCGCG 3575

Qy 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340

Db 3576 GCCATGCAAGCCCAACAGGACGATGAAGAAAGTTTTTCTCAGAGCGGGTTCTCATC 3635

Qy 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360

Db 3636 TTTGGGAAGCAAGGCTCAGAGAAAACAATGTGGACATTTGAAAGGTCATGATTACAGAC 3695

Qy 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380

Db 3696 GAAGAGGAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTATGGTTCTGTATCTACC 3755

Qy 381 AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400

Db 3756 AACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAACACACACAAGCGTTCTT 3815

Qy 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420

Db 3816 CCAGGCATGCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGCAAGATT 3875

Qy 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440

Db 3876 CCACACACGGACGGACATTTTCACCCCTTCCCTCATGGGTGGATTGGACTTAAACAC 3935

Qy 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460

Db 3936 CCTCCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAATCCTTCGACCACCTTC 3995

Qy 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480

Db 3996 AGTGGGCAAGTTTGCTTCTTCATCACACAGTACTCCACGGGACAGGTCAGCGTGGAG 4055

Qy 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 500

Db 4056 ATCGAGTGGGAGCTGCAGAAGGAAACAGCAACCGCTGGAATCCCGAAATTCAGTACACT 4115

Qy 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520

Db 4116 TCCAACACTACAACAAGTCTGTTAATGTGGACTTACTGTGGACACTAATGGCGTGTATTCA 4175

Qy 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534

Db 4176 GAGCCTCGCCCCCATGGCACCAGATACCTGACTCGTAATCTG 4217

RESULT 11

AR222045 LOCUS AR222045 7557 bp DNA linear PAT 26-SEP-2002

DEFINITION Sequence 3 from patent US 6429001.

ACCESSION AR222045

VERSION AR222045.1 GI:23329401

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 7557)

AUTHORS Hardy,S.F.

TITLE Recombinant AAV packaging systems

JOURNAL Patent: US 6429001-A 3 06-AUG-2002;

FEATURES Location/Qualifiers

source 1..7557

BASE COUNT 1970 a 1940 c 1953 g 1694 t

ORIGIN

Alignment Scores:

Pred. No.: 2.07e-191 Length: 7557

Score: 2481.50 Matches: 445

Percent Similarity: 90.26% Conservative: 37

Best Local Similarity: 83.33% Mismatches: 51

Query Match: 85.39% Indels: 1

DB: 6 Gaps: 1

US-09-807-802A-17 (1-534) x AR222045 (1-7557)

Qy 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20

Db 2837 ATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGGCGCCGCGAGTGGGT 2896

Qy 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40

Db 2897 AATTCTCGGGAATTGGCATTGCGATTCCACATGGATGGCGACAGAGTCATCACCACC 2956

Qy 41 SerThrArgThrTrpAlaLeuProThrTrpAsnAsnHisLeuTyrLysGlnIleSerSer 60

Db 2957 AGCACCAGAACCTGGGCCCTGCCACCTACAAACACCACTCTTACAAACAAATTTCCAGC 3016

Qy 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80

Db 3017 CAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTTGGGGGTAT 3073

Qy 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100

Db 3074 TTTGACTTCAACAGATTCCACTGCCACTTTTTCACCACGTGACTGGCAAGACTCATCAAC 3133

Qy 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120

Db 3134 AACAACTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAACATTCAAGTCAAA 3193

Qy 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140

Db 3194 GAGGTCACGCAGAAATGACGGTACGACGACGATTGCCAATAAACCTTACCAGCACCGGTTGAG 3253

Qy 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160

Db 3254 GTGTTACTGACTCGGAGTACACGCTCCCGTACGTCTCGGCTCGGCGCATCAAGGATGC 3313

Qy 161 LeuProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180

Db 3314 CTCCGCGCGTTCCAGCAGACGCTCTTCATGGTGGCCACAGTATGGATACCTCACCCCTGAAC 3373

Qy 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200

Db 3374 AACGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCTGGAGTACTTTTCCTTCTCAG 3433

Qy 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis 220

Db 3434 ATGCTGGGTACCGGAAACAACATTACCTTCAGCTACACTTTTGAGGACGTTCTCTTCCAC 3493

Qy 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240

Db 3494 AGCAGTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGACCATAC 3553

Qy 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260





Db 3614 TTTTCTCAGCGCGAGCGAGTGACATTTCGGGACCAGTCTAGGAACCTGGCTTCCTGGACCC 3673

Qy 281 CysTyrArgGlnGlnArgValSerLysThrLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
|||||  
Db 3674 TGTATCCGCCAGCAGCGAGTATCAAAGACATCTCGGGATAACAACAAGTGAATACCTCG 3733  
|||||

Qy 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320  
|||||  
Db 3734 TGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGCCCG 3793  
|||||

Qy 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340  
|||||  
Db 3794 GCCATGGCAAGCCACAGGACGATGAAGAAAGTTTTCCTCAGAGCGGGTTCATC 3853  
|||||

Qy 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360  
|||||  
Db 3854 TTTGGGAAGCAAGGCTCAGAGAAACAATGTGGACATTGAAAGGTTCATGATTACAGAC 3913  
|||||

Qy 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380  
|||||  
Db 3914 GAAGAGGAAATCAGGACAAACCAATCCCGTGGCTACGGAGCAGTATGGTTCTGTATCTACC 3973  
|||||

Qy 381 AsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400  
|||||  
Db 3974 AACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAACACACAGCGGTTCTT 4033  
|||||

Qy 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420  
|||||  
Db 4034 CCAGGCATGGTCTGGCAGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGCAAGATT 4093  
|||||

Qy 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440  
|||||  
Db 4094 CCACACACGGACGGACATTTTCACCCCTCTCCCCTCATGGTGGATTTCGGACTTAAACAC 4153  
|||||

Qy 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460  
|||||  
Db 4154 CCTCCTCCACAGATTCTCATCAAGAACACCCCGGTACTCGGAATCCTTCGACCACCTTC 4213  
|||||

Qy 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480  
|||||  
Db 4214 AGTGGGCAAGTTTGCTTCCTTCATCACACAGTACTCCACGGGACAGGTCAGCGTGGAG 4273  
|||||

Qy 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 500  
|||||  
Db 4274 ATCGAGTGGGAGCTGCAGAGGAAACAGCAACCGCTGGAATCCCGAAATTCAGTACACT 4333  
|||||

Qy 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520  
|||||  
Db 4334 TCCAACTACAACNAGTCTGTTAATGTGGACTTTTACTGTGGACACTAATGGCGGTATTCA 4393  
|||||

Qy 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
|||||  
Db 4394 GAGCCTCGCCCATTTGGCACCAGATACCTGACTCGTAACTCG 4435  
|||||

RESULT 13  
AR222044  
LOCUS AR222044 8698 bp DNA linear PAT 26-SEP-2002  
DEFINITION Sequence 2 from patent US 6429001.  
ACCESSION AR222044  
VERSION AR222044.1 GI:23329400  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 8698)  
AUTHORS Hardy,S.F.  
TITLE Recombinant AAV packaging systems  
JOURNAL Patent: US 6429001-A 2 06-AUG-2002;  
FEATURES Location/Qualifiers  
source 1..8698  
/organism="unknown"  
BASE COUNT 2117 a 2368 c 2306 g 1907 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2,49e-191 Length: 8698  
Score: 2481.50 Matches: 445  
Percent Similarity: 90.26% Conservative: 37  
Best Local Similarity: 83.33% Mismatches: 51  
Query Match: 85.39% Indels: 1  
DB: 6 Gaps: 1

US-09-807-802A-17 (1-534) x AR222044 (1-8698)

Qy 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20  
|||||  
Db 2809 ATGGCTACAGGCAGTGGCGCACCATGGCAGACAATAACGAGGCGCCGAGTGGGT 2868  
|||||

Qy 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40  
|||||  
Db 2869 AATTCCTCGGGAATTTGGCATTTGGCATTTCCACATGGATGGGCGACAGAGTATCACCACC 2928  
|||||

Qy 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
|||||  
Db 2929 AGCACCCGAACCTGGGCCCTGCCACCTTACAACAACACCTCTACAACAATAATTTCCAGC 2988  
|||||

Qy 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80  
|||||  
Db 2989 CAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGACCCCTTTGGGGGTAT 3045  
|||||

Qy 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100  
|||||  
Db 3046 TTTGACTTCAACAGATTCCACTGCCACTTTTCACCACGTGACTGGCAAGAGACTCATCAAC 3105  
|||||

Qy 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120  
|||||  
Db 3106 AACAACTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAAACATTTCAAGTCAA 3165  
|||||

Qy 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140  
|||||  
Db 3166 GAGGTACGCAGAAATGACGGTACGACGAGATTGCCAATAACCTTACCAGCAGCGTTTCAG 3225  
|||||

Qy 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160  
|||||  
Db 3226 GTGTTTACTGACTCGGAGTACCAGTCCCGTACGTCTCTCGGCTCGGCGCATCAAGGATGC 3285  
|||||

Qy 161 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
|||||  
Db 3286 CTCGCCCGCTTCCAGCAGACGCTCTCATGGTGCCACAGTATGGATACCTCACCTTGAAC 3345  
|||||

Qy 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
|||||  
Db 3346 AACGGGAGTCAGGCAGTAGGACGCTCTTCACTTTTACTGCCTGGAGTACTTCTCTCTCAG 3405  
|||||

Qy 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluValPropheHis 220  
|||||  
Db 3406 ATGCTGCGTACCGGAAACAACCTTTACTTTCAGCTACACTTTTGGAGGACGTTCTTTCCAC 3465  
|||||

Qy 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
|||||  
Db 3466 AGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCTCTCATCGACCAGTAC 3525  
|||||

Qy 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260  
|||||  
Db 3526 CTGTATTACTTGAGCAGAAACAACACTCCAAGTGGAAACCAACCGCAGTCAAGGCTTCAG 3585  
|||||

Qy 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280  
|||||  
Db 3586 TTTTCTCAGCGCGGAGCGAGTGACATTTCGGGACCAGTCTAGGAACCTGGCTTCCTGGACCC 3645  
|||||

Qy 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
|||||  
Db 3646 TGTACCGCCAGCAGCGAGTATCAAAAGACATCTGCGGATAACAACAACAGTGAATACTCG 3705  
|||||

Qy 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320  
|||||

|    |      |   |      |
|----|------|---|------|
| Db | 3706 | TGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGCCCG  | 3765 |
| QY | 321  | AlaMetAlaSerHisLysAspGluAspLysPhePheProMetSerGlyValMetIle     | 340  |
| Db | 3766 | GCCATGGCAAGCCACAAAGGACGATGAAGAAAGTTTTTTCCTCAGAGCGGGGTCTCATC   | 3825 |
| QY | 341  | PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp  | 360  |
| Db | 3826 | TTTGGGAAGCAAGGCTCAGAGAAAACAAATGTGGACATTGAAAAGGTTCATGATTACAGAC | 3885 |
| QY | 361  | GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal  | 380  |
| Db | 3886 | GAAGAGGAATCAGGACAAACCAATCCCGTGGCTACGGAGCAGTATGGTTCTGTATCTACC  | 3945 |
| QY | 381  | AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu  | 400  |
| Db | 3946 | AACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAACACACAAAGCGTTCTT  | 4005 |
| QY | 401  | ProGlyMetValTipGlnAspArgAspValTyrLeuGlnGlyProIleTipAlaLysIle  | 420  |
| Db | 4006 | CCAGGCATGGTCTGGCAGGACAGAGATGTACCTTCAGGGGCCCATCTGGGCAAGATT     | 4065 |
| QY | 421  | ProHisThrAspGlyHisPheHisProSerProLeuMetGlyPheGlyLeuLysAsn     | 440  |
| Db | 4066 | CCACACACGGACGGACATTTTCACCCCTCTCCCTCATGGGTGGATTCCGACCTTAAACAC  | 4125 |
| QY | 441  | ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe  | 460  |
| Db | 4126 | CCTCCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAATCCTTCGACCACCTTC  | 4185 |
| QY | 461  | SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu  | 480  |
| Db | 4186 | AGTGGGCAAGTTTGCTTCCTTCATCACACAGTACTCCACGGGACAGGTACGCGTGGAG    | 4245 |
| QY | 481  | IleGluTrpGluLeuGlnLysGluAsnSerLysArgTipAsnProGluValGlnTyrThr  | 500  |
| Db | 4246 | ATCGAGTGGGAGCTGCAGAGGAAAACAGCAACGCTGGAATCCCGAAATTCAGTACACT    | 4305 |
| QY | 501  | SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr  | 520  |
| Db | 4306 | TCCAACTACAACAAGTCTGTTAATGTGGACTTTACTGTGGACACTAATGGCGTGTATTCA  | 4365 |
| QY | 521  | GluProArgProIleGlyThrArgTyrLeuThrArgProLeu                    | 534  |
| Db | 4366 | GAGCCTCGCCCCATTGGCACCCAGATACCTGACTCGTAACTCG                   | 4407 |

RESULT 14  
AX205072 LOCUS AX205072 8698 bp DNA linear PAT 30-AUG-2001  
DEFINITION Sequence 2 from Patent WO0155361.  
ACCESSION AX205072  
VERSION AX205072.1 GI:15394313

KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified.  
REFERENCE 1  
AUTHORS Hardy,S.F.  
TITLE Recombinant aav packaging systems  
JOURNAL Patent: WO 0155361-A 2 02-AUG-2001;  
Chiron Corporation (US)  
FEATURES  
source Location/Qualifiers  
1..8698  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"  
/note="recombinant DNA"

BASE COUNT 2117 a 2368 c 2306 g 1907 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.49e-191 Length: 8698  
Score: 2481.50 Matches: 445

|  |        |  |      |
|--|--------|--|------|
| Percent Similarity:                            | 90.26% | Conservative:  | 37   |
| Best Local Similarity:                         | 83.33% | Mismatches:  | 51   |
| Query Match:                                   | 85.39% | Indels:  | 1    |
| DB:  | 6      | Gaps:  | 1    |
| US-09-807-802A-17' (1-534) x AX205072 (1-8698) |        |  |      |
| QY   | 1      | MetAlaSerGlyGlyAlaProMetAlaAspAsnGluGlyAlaAspGlyValGly         | 20   |
| Db   | 2809   | ATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGGCGCCGCGAGTGGGT     | 2868 |
| QY   | 21     | AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr   | 40   |
| Db   | 2869   | AATTCTCGGGAAATTGGCATTGCGATTCCACATGGATGGCGACAGAGTCATCACCACC     | 2928 |
| QY   | 41     | SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer   | 60   |
| Db   | 2929   | AGCACCCGAACCTGGGGCCCTGCCACCTACAACAACCAACCTCTACAAAACAATTTCCAGC  | 2988 |
| QY   | 61     | AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr   | 80   |
| Db   | 2989   | CAATCA--GGAGCCTCGAACGACAATCATACTATTGGCTACAGCACCCCTTGGGGGTAT    | 3045 |
| QY   | 81     | PheAspPheAsnArgPheHisCysHisPheSerProArgAspTipGlnArgLeuIleAsn   | 100  |
| Db   | 3046   | TTTGACTTCAACAGATTCCACTGCCACTTTTCAACACGTGACTGGCAAAGACTCATCAAC   | 3105 |
| QY   | 101    | AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys   | 120  |
| Db   | 3106   | AACAACCTGGGGATTCGACCCCAAGAGACTCAACTTCAAGCTCTTTAAACATTCAAGTCAA  | 3165 |
| QY   | 121    | GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln   | 140  |
| Db   | 3166   | GAGGTACGCGAATGACGGGTACGACGACGATTGCCAATAACCTTACCAGCACGCGTTGAG   | 3225 |
| QY   | 141    | ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys   | 160  |
| Db   | 3226   | GTGTTTACTGACTCGGAGTACCAGTCCCGTACGTCTCGGCTCGGCGCATCAAGGATGC     | 3285 |
| QY   | 161    | LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn   | 180  |
| Db   | 3286   | CTCCCGCGCTTCCAGCAGACGCTTTCATGGTGCCACAGTATGGATACCTCACCCCTGAAC   | 3345 |
| QY   | 181    | AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln   | 200  |
| Db   | 3346   | AACGGGAGTCAGGCAGTAGGACGCTCTTTCATTTTACTGCTGGAGTACTTCTCTTCTCAG   | 3405 |
| QY   | 201    | MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis   | 220  |
| Db   | 3406   | ATGCTGCGTACCGGAAACAACCTTTACCTTCAGTACACCTTTTGAGGACGTTCTTCTCCAC  | 3465 |
| QY   | 221    | SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr   | 240  |
| Db   | 3466   | AGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGACCAGTAC   | 3525 |
| QY   | 241    | LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu   | 260  |
| Db   | 3526   | CTGTATTACTTGAGCAGAAACAACACTCCAAGTGGAAACCAACACGCGAGTCAAGGCTTCAG | 3585 |
| QY   | 261    | PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro   | 280  |
| Db   | 3586   | TTTTTCTCAGGCGGAGCGAGTGACATTTCGGGACCAAGTCTAGGAACCTGGCTTCTGGACCC | 3645 |
| QY   | 281    | CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr   | 300  |
| Db   | 3646   | TGTTACCGCCAGCAGGAGTATCAAAGACATCTCGGGATACAACAACAGTGAATACTCG     | 3705 |
| QY   | 301    | TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr   | 320  |
| Db   | 3706   | TGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGTTGAATCCGGGCCCCG  | 3765 |
| QY   | 321    | AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle   | 340  |

Db 3766 GCCATGGCAAGCCACAAGGACGATGAAGAAAAGTTTTCCTCAGACGGGGTTCTCATC 3825

Qy 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360

Db 3826 TTTGGGAAGCAAGGCTCAGAGAAAACAATGTGGACATTGAAAGGTCATGATTACAGAC 3885

Qy 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380

Db 3886 GAAGAGGAAATCAGGACAACCAATCCCCTGCTACGGAGCAGTATGTTCTGTATCTACC 3945

Qy 381 AsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400

Db 3946 AACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAACACACAAGGCGTTCTT 4005

Qy 401 ProGlyMetValTrpGlnAspArgAspValTrpLeuGlnGlyProIleTrpAlaLysIle 420

Db 4006 CCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGCAAGATT 4065

Qy 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440

Db 4066 CCACACACGGACGGACATTTTCACCCCTCTCCCTCATGGGTGGATTCCGACTTAAACAC 4125

Qy 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460

Db 4126 CCTCCTCCACAGATTCTCATCAAGAACACCCCGGTACCTCGGAATCCTTCGACCACCTTC 4185

Qy 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480

Db 4186 AGTGGGCAAGTTTGGTCTCTTCATCACACAGTACTCCACGGGACAGTCCAGCGTGGAG 4245

Qy 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 500

Db 4246 ATCGAGTGGGAGCTGCAGAAGGAAACAGCAACCGTGGAAATCCCGAAATTCAGTACACT 4305

Qy 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520

Db 4306 TCCAACACTACAACAAGTCTGTTAATGTGGACTTTACTGTGGACACTAATGGCGTGATTCA 4365

Qy 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534

Db 4366 GAGCCTCGCCCATTTGGCACCAGATACCTGTACTCGTAATCTG 4407

RESULT 15

AY243021

LOCUS

DEFINITION Non-human primate Adeno-associated virus isolate AAVch.5 capsid protein (VP1) gene, complete cds.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. .2208

/organism="Non-human primate Adeno-associated virus"

/proviral

/mol\_type="genomic DNA"

/isolate="AAVch.5"

/db\_xref="taxon:226582"

1. .2208

/gene="VP1"

1. .2208

/gene="VP1"

/codon\_start=1

/product="capsid protein"

/protein\_id="AAO88207.1"

/db\_xref="GI:29650538"

/translation="MAADGYLPDWLEDTLSEGIQRMWKLKPGPPPPKPNQHRDDSRG

LVLPGYKYLGFNGLDKGEPVNEADAALEHDKAYDHLKQGDNPYLKYNHADAEPQBE

RLQEDTSPGNGLGRAVFOAKQVLEPLGLVEBAVKTAPGKRPVIEQSPAEPDSSGIG

KSGQQPAKKRLNFGQTDTESVDPDQPIGEPPAAPSGVGSNTMASGGGAPMADNNEGA

DGVNSSGNWHCDSTWMDRVITSTRTWALPTYNHLYKQISSSESGATNDNHYPGYS

TPWGYFDNRFHCHFSRDRWQRLINNNWGFPRKLNFKLFNIQVKEVTQNDGTTIAN

NLTSTVQVFTDSEYQLPYVLGSAHQCLPPPADVFMIPOYGYLTLLNGSOAVGRSSP

YCLEYFPPSQMLRTGNNTFTSYTFEDVPFHSSVAHSQSLDRMLNPLIDQLYLYLSKTQG

TSGTTQQSRLQFSQAGPSSMAQQAQKWLPGSPYRQQRMSKTANDNNSEFAPWTAATKY

YLNGRNSLVNPGPMASHKDEEKYFPMHGNIIFGKQGTGTNNVDIESVLITDEEIR

TTNPVATEQYQVATNHQSQNTTASYSQSVDSQILPGMWQDRDVLQGPWAKTPHT

DGHFHPSPLMGGFGLKHPPIILIKNTVPANPATFTPGKFASFITQYSTGQVSVEI

EWELQKENSKRWNPEIQYTSYNKSVNVEFTVDANGVYSEPRPIGTRYLTRNL"

BASE COUNT 617 a 638 c 518 g 435 t

ORIGIN

Alignment Scores:

Pred. No.: 3.6e-190 Length: 2208

Score: 2457.50 Matches: 442

Percent Similarity: 89.70% Conservative: 37

Best Local Similarity: 82.77% Mismatches: 54

Query Match: 84.57% Indels: 1

DB: 14 Gaps: 1

US-09-807-802A-17 (1-534) x AY243021 (1-2208)

Qy 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20

Db 607 ATGGCTTCAGGCGGTGGGGCACCAATGGCAGACAATAACGAAGGCCCGCAGGTGGGT 666

Qy 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40

Db 667 AATTCCTCGGGAATTTGGCATTTCCATTCACATGGATGGGCGCAGAGTTATCACCACC 726

Qy 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60

Db 727 AGCACAGAAACCTGGGCCCTCCCCACCTACAATAATACCTCTACAAGCAATCTCCAGC 786

Qy 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80

Db 787 GAATCG--GGAGCCACCACCAACGACAAACCACTACTTTCGGCTACAGACCCCTCGGGGTAT 843

Qy 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100

Db 844 TTTGACTTTAAACAGATTCCACTGTCACTTCTCACCACGTGACTGGCAGCGACTCATCAAC 903

Qy 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120

Db 904 AACAACTGGGGATTAGACCCCAAGAAACTCAATTTCAAGCTCTTCAACATCCAAGTCAAG 963

Qy 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140

Db 964 GAGGTACCGCAGAATGATGGAAACCAACGACCATCGCCCAATAACCTTACCAGACCGGTGCAG 1023

Qy 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160

Db 1024 GTCTTCACAGACTCTGAGTACCAGCTGCGCTACGTCTCGGTTTCGGCTCACCAGGGCTGC 1083

Qy 161 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180

Db 1084 CTTCCGCGGTTCCCGACGACGACGTCTTCATGATTCCTCAGTACGGCTACTTGACTCTGAAC 1143

Qy 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200



Db 1144 AATGGCAGCCAAGCGGTAGGACGTTCTTCTACTGTCTAGAGTATTTTCCCTCTCAG 1203

Qy 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPropPheHis 220  
|||||  
|||||

Db 1204 ATGCTGAGGACGGGAAACAATTCACCTTCAGCTACACTTTTGAAGACGTGCTTCCAC 1263

Qy 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
|||||  
|||||

Db 1264 AGCAGCTACGCGCACAGCCAGAGTCTGGATCGGCTGATGAATCCTCTCATTGACCAGTAC 1323

Qy 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260  
|||||  
|||||

Db 1324 CTGTATTACCTGAGCAAAACTCAGGGGTACAAGTGGAAACAACGACGCAATCGAGACTGCAG 1383

Qy 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTyrLeuProGlyPro 280  
|||||  
|||||

Db 1384 TTCAGCCAAGCTGGGCCTAGTCCATGGCTCAGCAGGCCAAAAACTGGCTACCGGGACCC 1443

Qy 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
|||||  
|||||

Db 1444 AGCTACCGACAGCAGCGAATGTCTAAGACGGCTAATGACAACAACACAGTGAATTGCT 1503

Qy 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320  
|||||  
|||||

Db 1504 TGGACTGCAGCCACCAATATTACCTGAATGGAAGAAATTCTCTGGTCAATCCCGGCCCC 1563

Qy 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340  
|||||  
|||||

Db 1564 CCAATGGCCAGTCACAAGGACGATGAGGAAAAAGTATTTCCCATGCACGGAAATCTCATC 1623

Qy 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360  
|||||  
|||||

Db 1624 TTTGAAAAACAAGGCACAGGAACCTACCAATGTGGACATTGAATCAGTGTCTATTACAGAC 1683

Qy 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380  
|||||  
|||||

Db 1684 GAAGAAGAAATCAGAACAACTAATCCTGTGGCTACAGAACAAATACGGACAGGTTGCCACC 1743

Qy 381 AsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400  
|||||  
|||||

Db 1744 AACCATCAGAGTCAGAACACCCACAGCTTCCTATGGAAGTGTGGACAGCCAGGGGAATCTTA 1803

Qy 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaIysIle 420  
|||||  
|||||

Db 1804 CCTGGAATGGTGTGGCAGGACCGCGATGTCTATCTTCAAGGTCCCATTGGGCCAAAACT 1863

Qy 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440  
|||||  
|||||

Db 1864 CCTCACGGGACGGACACTTTTCATCCTTCTCCGCTCATGGGAGGCTTTGGACTGAAACAC 1923

Qy 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460  
|||||  
|||||

Db 1924 CCTCCTCCCCAGATCCTGTATCAAAAAACACACCTGTGCCAGGAATCCCGCGACCACTTTC 1983

Qy 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480  
|||||  
|||||

Db 1984 ACTCCTGGAAAGTTTGCTTCGTTTCATTACCCAGTATTCACCGGACAGGTGAGCTGGAA 2043

Qy 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTyrAsnProGluValGlnTyrThr 500  
|||||  
|||||

Db 2044 ATAGAGTGGGAGCTGCAGAAAGAAACAGCAACCGCTGGAACCCAGAAATTCAGTACACC 2103

Qy 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520  
|||||  
|||||

Db 2104 TCCAACTACAAACAAGTCGGTGAATGTGGAGTTTACCCTGGACGCAACCGTGTATTCT 2163

Qy 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
|||||  
|||||

Db 2164 GAACCCCGCCCTATTGGCACTCGTTACCTTACCCCGGAACCTG 2205

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 14, 2004, 15:18:00 ; Search time 366.286 Seconds  
(without alignments)  
3935.452 Million cell updates/sec

Title: US-09-807-802A-17  
Perfect score: 2906  
Sequence: 1 MASGGGAPMADNNEGADGVG.....NNGLYTEPRPIGTRYLTRPL 534

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US9807802/runat\_11022004\_175606\_15919/app\_query.fasta\_1.2389  
-DB=N Geneseq\_19Jun03 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosom62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US9807802 @CGN 1 1 806 @runat\_11022004\_175606\_15919 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq\_19Jun03:\*  
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*  
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Match | Length | DB | ID        | Description        |
|------------|--------|-------|--------|----|-----------|--------------------|
| 1          | 2906   | 100.0 | 1605   | 21 | AAD00779  | Adeno-associated v |
| 2          | 2906   | 100.0 | 1800   | 21 | AAD00778  | Adeno-associated v |
| 3          | 2906   | 100.0 | 2211   | 21 | AAD00777  | Adeno-associated v |
| 4          | 2906   | 100.0 | 4718   | 21 | AAD00772  | Adeno-associated v |
| 5          | 2884   | 99.2  | 4683   | 22 | AAF233749 | AAV6 DNA sequence. |
| 6          | 2530   | 87.1  | 4722   | 22 | AAF233748 | AAV3B DNA sequence |
| 7          | 2518   | 86.6  | 8178   | 21 | AAD00834  | Chimeric adeno-ass |
| 8          | 2481.5 | 85.4  | 4072   | 22 | AAH26327  | Adenovirus helper  |
| 9          | 2481.5 | 85.4  | 4679   | 22 | AAI66974  | Adeno-associated v |
| 10         | 2481.5 | 85.4  | 4679   | 22 | AAF233750 | AAV2 DNA sequence. |
| 11         | 2481.5 | 85.4  | 4679   | 24 | ABK89694  | Adeno-associated v |
| 12         | 2481.5 | 85.4  | 4679   | 25 | ABV76133  | Adeno associated v |
| 13         | 2481.5 | 85.4  | 7557   | 22 | AAH26326  | Adeno-associated v |
| 14         | 2481.5 | 85.4  | 8698   | 22 | AAH26325  | Wild-type adeno-as |
| 15         | 2439.5 | 83.9  | 4680   | 17 | AAT09008  | Wild-type adeno-as |
| 16         | 2439.5 | 83.9  | 4680   | 25 | ABX14497  | Wild-type adeno-as |
| 17         | 2423.5 | 83.4  | 4675   | 22 | AAH41481  | Adeno-associated v |
| 18         | 2423.5 | 83.4  | 4675   | 22 | AAF89931  | Nucleotide sequenc |
| 19         | 2423.5 | 83.4  | 4675   | 24 | ABS69879  | Human adeno-associ |
| 20         | 2423.5 | 83.4  | 4675   | 24 | ABS69880  | Human adeno-associ |
| 21         | 2423.5 | 83.4  | 4675   | 24 | ABA02989  | Adeno-associated v |
| 22         | 2258.5 | 77.7  | 8151   | 21 | AAD00832  | Chimeric adeno-ass |
| 23         | 1690.5 | 58.2  | 1617   | 19 | AAV21658  | AAV4 VP3 coat prot |
| 24         | 1690.5 | 58.2  | 1617   | 25 | ABX15817  | Adeno-associated v |
| 25         | 1690.5 | 58.2  | 1800   | 19 | AAV21657  | AAV4 VP2 coat prot |
| 26         | 1690.5 | 58.2  | 1800   | 25 | ABX15816  | Adeno-associated v |
| 27         | 1690.5 | 58.2  | 2208   | 19 | AAV21650  | AAV4 VP1 capsid pr |
| 28         | 1690.5 | 58.2  | 2208   | 25 | ABX15809  | Adeno-associated v |
| 29         | 1690.5 | 58.2  | 4767   | 19 | AAV21648  | AAV4 genome. Aden  |
| 30         | 1690.5 | 58.2  | 4767   | 25 | ABX15807  | Adeno-associated v |
| 31         | 1690   | 58.2  | 2487   | 18 | AAT69854  | Duck parvovirus ca |
| 32         | 1680.5 | 57.8  | 7215   | 21 | AAD00831  | Hybrid adeno-assoc |
| 33         | 1676   | 57.7  | 4570   | 16 | AAT01190  | Barbary duck parvo |
| 34         | 1665   | 57.3  | 2264   | 21 | AAZ49213  | Adeno associated v |
| 35         | 1665   | 57.3  | 2264   | 21 | AAZ49214  | Adeno associated v |
| 36         | 1665   | 57.3  | 2307   | 21 | AAZ49212  | Adeno associated v |
| 37         | 1665   | 57.3  | 4652   | 21 | AAZ49209  | Adeno associated v |
| 38         | 1665   | 57.3  | 4652   | 24 | AAS17712  | Adebo-associated v |
| 39         | 491.5  | 16.9  | 2271   | 21 | AAD00833  | Adeno-associated v |
| 40         | 490    | 16.9  | 1662   | 20 | AAH81586  | Erythrovirus V9 DN |
| 41         | 490    | 16.9  | 2343   | 20 | AAH81583  | Erythrovirus V9 DN |
| 42         | 490    | 16.9  | 4677   | 16 | AAT49535  | Human parvovirus g |
| 43         | 490    | 16.9  | 5028   | 20 | AAH81580  | Genomic DNA sequen |
| 44         | 486    | 16.7  | 1699   | 25 | ABZ59574  | Human parvovirus B |
| 45         | 486    | 16.7  | 1699   | 25 | ABZ59577  | Human parvovirus B |

ALIGNMENTS

RESULT 1  
AAD00779  
ID AAD00779 standard; DNA; 1605 BP.  
XX  
AC AAD00779;  
XX  
DT 08-SEP-2000 (first entry)  
XX  
DE Adeno-associated virus serotype 1 capsid protein VP3 DNA.  
KW Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;  
KW cap protein; recombinant viral vector; gene delivery; gene therapy;  
KW vaccine; transgene; VP3; ss.  
XX  
OS Adeno associated virus serotype 1.  
XX  
FH Location/Qualifiers

```
FT CDS 1..1605
FT /*tag= a
FT /product= "VP3 protein"
XX WO200028061-A2.
XX
XX 18-MAY-2000.
XX
XX 02-NOV-1999; 99WO-US25694.
XX
XX 05-NOV-1998; 98US-0107114.
XX
XX (UYPE-) UNIV PENNSYLVANIA.
XX
XX Wilson JM, Xiao W;
XX
XX WPI; 2000-376571/32.
XX
XX P-PSDB; AAY71169.
XX
XX Novel adeno-associated virus serotype 1 polynucleotide useful for
XX preparation of medicament for delivery of a transgene to a host
XX
XX Claim 10; Page 96-99; 108pp; English.
XX
XX The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA
XX which is characterised by two inverted terminal repeats (ITR) and open
XX reading frames for rep and capsid (cap) proteins. The rep reading frame
XX encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap
XX reading frame encodes three structural proteins, VP1, VP2 and VP3.
XX The AAV-1 sequence or its fragments particularly ITRs, rep and cap
XX coding regions, are useful in production of recombinant viral vectors
XX for gene delivery. These vectors can be used as gene therapy
XX vectors, vaccine vectors or antisense delivery vectors. The AAV-1 does
XX not induce the formation of neutralising antibodies specific to any
XX serotype of AAV hence is useful for transforming host cells, and in the
XX preparation of a medicament for the delivery of transgene to a host.
XX The present sequence is an AAV-1 DNA encoding a cap protein VP3 which is
XX useful in the production of recombinant viral vector for gene delivery.
XX
XX Sequence 1605 BP; 431 A; 466 C; 361 G; 347 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 1.07e-268 Length: 1605
Score: 2906.00 Matches: 534
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0
US-09-807-802A-17 (1-534) x AAD00779 (1-1605)
QY 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20
Db 1 ATGGCTTCAGGCGGTGGCCACCAATGGCAGACAAATAACGAAGGCCCGCGAGTGGGT 60
QY 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40
Db 61 AATGCCTCAGGAAATTGGCAATTGGATTCCACATGGCTGGCGGACAGAGTCAACCACC 120
QY 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60
Db 121 AGCACCCGCACCTGGGCTTGGCCACCTACATAACACCCTCTACAAGCAATCTCCAGT 180
QY 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80
Db 181 GCTTCAACGGGGGCCAGCAACGACAAACCACTACTTCGGCTACAGCACCCCTGGGGGTAT 240
QY 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100
Db 241 TTGATTTCAACAGATTCCACTGCCACTTTTCACCACGTGACTGGCAGCGACTCATCAAC 300
QY 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPhelysLeuPheAsnIleGlnValLys 120
Db
```

```
Db 301 AACAAATTGGGGATTCCGGCCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAGTCAAG 360
QY 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140
Db 361 GAGGTCAACGACGATGATGGGTCAACAACCATCGCTAATAAATACCTTACCAGCACGGTTCAA 420
QY 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160
Db 421 GTCTTCTCGGACTCGGAGTACCAGCTTCCGTACGTCTCGGCTCTGCGCACCGGCTGC 480
QY 161 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180
Db 481 CTCCCTCCGTTCCCGCGGACGTGTTCATGATTCGGAATACGGCTACCTGACGCTCAAC 540
QY 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200
Db 541 AATGGCAGCCCAAGCCGTGGGACGTTTCATCCTTTTACTGCTGGAATATTTCCCTTCTCAG 600
QY 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis 220
Db 601 ATGCTGAGAACGGGCAACAACCTTACCTTACGTACACCTTTGAGGAAGTGCCTTTCCAC 660
QY 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240
Db 661 AGCAGTACGCGCACAGCAGAGCCTGGACCGCTGATGAATCCTCTCATCGACCAATAC 720
QY 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260
Db 721 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGGAAGTGCCCAAAACAAGGACTTGTG 780
QY 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280
Db 781 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAACTGGCTACCTGGACCC 840
QY 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300
Db 841 TGTATCGGCAGCAGCGCGGTCTTCTAAACAACAAACAGACAAACAACAGCAATTTTACC 900
QY 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320
Db 901 TGGACTGGTGTCTTCAAAATATAAACCTCAATGGCGTGAATCCATCATCAACCCCTGGCACT 960
QY 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340
Db 961 GCTATGGCCTCACACAAGACGACGAAGACAAGTTCTTTCCCATGAGCGGTGTGATGATT 1020
QY 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360
Db 1021 TTTGGAAAAGAGAGCGCGCGGAGCTTCAAAACACTGCAATTGGACAATGTGATGATTACAGAC 1080
QY 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380
Db 1081 GAAGAGGAAATTAAGCCACTTAACCTGTGGCCACCCGAAGAATTTGGACCGGTGGCAGTC 1140
QY 381 AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400
Db 1141 AATTTCCAGAGCAGCAGCAGACACCCTGCGACCCGGAGATGTGATGCTATGGGAGCATTA 1200
QY 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420
Db 1201 CCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCAGGTGCCATTTGGGCCAAATTT 1260
QY 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440
Db 1261 CCTCACACAGATGGACACTTTACCCCGTCTCCTCTTATGGCGGCTTTGGACTCAAGAAC 1320
QY 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460
Db 1321 CCGCCTCTCAGATCCTCATCAAAAACACGCGCTGTTCTCGGAATCCTCCGGCGGAGTTT 1380
QY 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480
Db 1381 TCAGCTACAAGTTTGTCTTCATTCATCAACCAATACTCCACAGGACAAGTGAGTGTGGAA 1440
```





```
Qy 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340
Db 1156 GCTATGGCCTCACACAAAGACGACGAAGACAAGTTCTTTCCCATGAGCGGTGTCTATGATT 1215

Qy 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360
Db 1216 TTTGGAAAAGAGAGCGCGGAGCTTCAACACTGCTGATTGGACAATGTCTATGATTACAGAC 1275

Qy 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380
Db 1276 GAAGAGGAAATTAAAGCCACTAACCCCTGTGGCCACCAGAAAGATTGGGACCGTGGCAGTC 1335

Qy 381 AsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400
Db 1336 AATTTCCAGAGCAGCAGCACAGACCCCTGCGACCGGAGATGTGTCATGCTATGGGAGCAATTA 1395

Qy 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420
Db 1396 CCTGGCATGGTGTGGCAAGATAGAGACGCTGTACCTGCAGGCTCCCATTTGGGCCAAAATT 1455

Qy 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440
Db 1456 CCTCACACAGATGGACACTTTTCACCCGCTCTCCTCTTATGGGCGGCTTTGGACTCAAGAAC 1515

Qy 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460
Db 1516 CCGCCTCTCAGATCCTCATCAAAAACACGCTGTTCTCTCGAATCCTCCGGCGGAGTTT 1575

Qy 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480
Db 1576 TCAGCTACAAAGTTTGTCTTCTTTCATTCACCCCAATACTCCACAGCAAGTGTGGAA 1635

Qy 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 500
Db 1636 ATTGAATGGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 1695

Qy 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520
Db 1696 TCCAATTATGCAAAAATCTGCCAACGTTGATTTTACTGTGGACAAACAATGGACTTTTATACT 1755

Qy 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534
Db 1756 GAGCCTCGCCCCCATTTGGCACCCGTTACCTTACCCGTCCTCCCTG 1797

RESULT 3
AAD00777
ID AAD00777 standard; DNA; 2211 BP.
XX
AC AAD00777;
XX
DT 08-SEP-2000 (first entry)
XX
DE Adeno-associated virus serotype 1 capsid protein VP1 DNA.
XX
KW Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;
KW cap protein; recombinant viral vector; gene delivery; gene therapy;
KW vaccine; transgene; VP1; ss.
XX
OS Adeno associated virus serotype 1.
XX
FH Key Location/Qualifiers
FT CDS 1..2211
FT /tag= a
FT /product= "VP1 protein"
XX
PN WO200028061-A2.
XX
PD 18-MAY-2000.
XX
PF 02-NOV-1999; 99WO-US25694.
XX
PR 05-NOV-1998; 98US-0107114.
XX
```

```
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Wilson JM, Xiao W;
XX
DR WPI; 2000-376571/32.
XX P-PSDB; AAY71167.
PT Novel adeno-associated virus serotype 1 polynucleotide useful for
XX preparation of medicament for delivery of a transgene to a host
PS Claim 10; Page 83-87; 108pp; English.
XX
CC The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA
CC which is characterised by two inverted terminal repeats (ITR) and open
CC reading frames for rep and capsid (cap) proteins. The rep reading frame
CC encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap
CC reading frame encodes three structural proteins, VP1, VP2 and VP3.
CC The AAV-1 sequence or its fragments particularly ITRs, rep and cap
CC coding regions, are useful in production of recombinant viral vectors
CC for gene delivery. These vectors can be used as gene therapy
CC vectors, vaccine vectors or antisense delivery vectors. The AAV-1 does
CC not induce the formation of neutralising antibodies specific to any
CC serotype of AAV hence is useful for transforming host cells, and in the
CC preparation of a medicament for the delivery of transgene to a host.
CC The present sequence is an AAV-1 DNA encoding a cap protein VP1 which is
CC useful in the production of recombinant viral vector for gene delivery.
XX
SQ Sequence 2211 BP; 577 A; 652 C; 538 G; 444 T; 0 other;

Alignment Scores:
Pred. No.: 1.7e-268 Length: 2211
Score: 2906.00 Matches: 534
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

US-09-807-802A-17 (1-534) x AAD00777 (1-2211)

Qy 1 MetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20
Db 607 ATGGCTTCAGCGGTGGCGACCAATGGCAGACATATACGAAGGCGCCGACGGAGTGGGT 666

Qy 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40
Db 667 AATGCCTCAGGAAATTTGGCATTTCCGATTCCACATGCTGGCGGACAGAGTCAATCCACC 726

Qy 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60
Db 727 AGCACCCGCACCTGGGCTTGGCCACCTACATAACCACTCTTACAGCAATCTCCAGT 786

Qy 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80
Db 787 GCTTCAACGGGGGCCAGCAACGACCAACCACTACTTTCGGCTACAGCACCCCTGGGGGTAT 846

Qy 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100
Db 847 TTTGATTTCAACAGATTCCCACTGCCACTTTTTCACCACGTCGACTGGCAGCGACTCATCAAC 906

Qy 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120
Db 907 AACAAATGGGGATTTCGGGCCCAAGAGACTCAACTTCAAACTTTTCAACATCCCAAGTCAAG 966

Qy 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140
Db 967 GAGGTCACGACGAATGATGGCGTCAACAACCATCGCTAATAACCTTACCAGCACCGGTTCAA 1026

Qy 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160
Db 1027 GTCTTCTCGGACTCGGAGTACCAGCTTCCGTACGTCCTCGGCTCTGCGCACCGGGCTGC 1086

Qy 161 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180
```

Db 1087 CTCCTCCGTTCCCGCGGACGTGTTTCATGATTCGCCAATACGGGTACCTGACGCTCAAC 1146

Qy 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
|||||

Db 1147 AATGGCAGCCAGCCGCGGACGTTCATCCTTTTACTGCTGGAATATTTCCCTTCTCAG 1206

Qy 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis 220  
|||||

Db 1207 ATGCTGAGAACGGGCAACAACCTTACCTTCAGCTACACCTTGAGGAAGTGCCTTTCCAC 1266

Qy 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
|||||

Db 1267 AGCAGCTACGCGCACAGCCAGAGCCTGGACCGGTGATGAATCCTCTCATCGACCAATAC 1326

Qy 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260  
|||||

Db 1327 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCTCAACAAAGGACTTGCTG 1386

Qy 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280  
|||||

Db 1387 TTTAGCCGTGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAACTGGCTACCTGGACCC 1446

Qy 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
|||||

Db 1447 TGTATCGGCAGCAGCGCGTTTCTAAACAAACAAACAGACAAACAGCAATTTTACC 1506

Qy 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320  
|||||

Db 1507 TGGACTGGTGTCTTCAAAATATTAACCTCAATGGCGGTGATCCATCATCAACCCCTGGCACT 1566

Qy 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340  
|||||

Db 1567 GCTATGCGCTCACACAAAGACGACGAAAGACAAGTTCTTCCCATGAGCGGTGTCAATGATT 1626

Qy 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360  
|||||

Db 1627 TTTGGAAAAGAGAGCGCGGAGCTTCAACACACTGCATTGGACAATGTCTATGATTACAGAC 1686

Qy 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380  
|||||

Db 1687 GAAGAGAAATTAAGCCACTAACCTGTGGCCACCAGAAAGATTGGGACCGTGGCAGTTC 1746

Qy 381 AsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400  
|||||

Db 1747 AATTTCCAGACGACGACGACGACCCCTGCGACCGGAGATGTGCATGCTATGGAGCATTA 1806

Qy 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420  
|||||

Db 1807 CCTGGCATGTGTGGCAAGATAGAGACGTGTACCTGCAGGTGCCATTTGGGCCAAAATT 1866

Qy 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440  
|||||

Db 1867 CCTCACACAGATGGACACTTTCACCCGTCTCTCTTATGGCGGCTTTGGACTCAAGAAC 1926

Qy 441 ProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460  
|||||

Db 1927 CCGCCTCCTCAGATCCTCATCAAAACACGCTGTTCTCTGCGAATCCTCCGCGGAGTTT 1986

Qy 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480  
|||||

Db 1987 TCAGCTACAAAGTTTGCTTTCATTCACCAATACTCCACAGACAAGTGAGTGGAA 2046

Qy 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 500  
|||||

Db 2047 ATTGAATGGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 2106

Qy 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520  
|||||

Db 2107 TCCAATTATGAAAAATCTGCCAACGTTGATTACTGTGGACAACAATGGACTTTTATACT 2166

Qy 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
|||||

Db 2167 GAGCCTCGCCCATTTGGCACCCCGTTTACCTTACCCGTCCCTG 2208

RESULT 4

AAD00772

ID AAD00772 standard; DNA; 4718 BP.

XX

AC AAD00772;

XX 08-SEP-2000 (first entry)

XX

DE Adeno-associated virus serotype 1 DNA.

XX

KW Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein; cap protein; recombinant viral vector; gene delivery; gene therapy; vaccine; transgene; ss.

KW

XX

OS Adeno associated virus serotype 1.

XX

FH Key Location/Qualifiers

FT repeat\_unit 1..143

FT /tag= a

FT /label= 5' ITR

FT /note= "Inverted terminal repeat which is capable of forming T-shaped hairpin structure"

FT protein\_bind 89..110

FT /tag= b

FT /bound moiety= "Rep protein"

FT misc\_feature 124..125

FT /tag= c

FT /note= "Terminal resolute site (TRS)"

FT misc\_binding 219..226

FT /tag= d

FT /bound moiety= "USP"

FT /note= "E box"

FT promoter 236..299

FT /tag= e

FT /label= P5\_promoter

FT protein\_bind 237..245

FT /tag= f

FT /bound moiety= "YY1 factor"

FT TATA\_signal 270..275

FT /tag= g

FT /label= P5\_TATA-Box

FT misc\_feature 299..306

FT /tag= h

FT /note= "YY1/p5 RNA"

FT CDS 335..2206

FT /tag= i

FT /product= "Rep 78"

FT /function= "regulates replication and integration of AAV DNA into host cell's chromosome"

FT CDS 335..2272

FT /tag= j

FT /product= "Rep 68"

FT /function= "regulates replication and integration of AAV DNA into host cell's chromosome"

FT /note= "The coding region is interrupted by intron"

FT intron 1924..2220

FT /tag= k

FT /note= "This region interrupts the coding sequence of Rep 68 and Rep 40"

FT TATA\_signal 857..862

FT /tag= l

FT /label= P19\_TATA\_Box

FT misc\_feature 882..883

FT /tag= m

FT /note= "P19 RNA"

FT CDS 1007..2206

FT /tag= n

FT /product= "Rep 52"

FT /function= "regulates replication and integration of AAV DNA into host cell's chromosome"

FT CDS 1007..2272

FT /tag= o



```
FT /product= "Rep 40"
FT /function= "regulates replication and integration
FT of AAV DNA into host cell's chromosome"
FT /note= "The coding region is interrupted by intron"
FT 1836..1841
FT /*tag= p
FT /label= P40_TATA-BOX
FT 1875..1876
FT /*tag= q
FT /note= "P40 RNA"
FT 2223..4433
FT /*tag= r
FT /product= "VP1 protein"
FT /note= "Capsid protein"
FT 2634..4433
FT /*tag= s
FT /product= "VP2 protein"
FT /note= "Capsid protein"
FT /partial
FT 2829..4433
FT /*tag= t
FT /product= "VP3 protein"
FT /note= "Capsid protein"
FT 4447..4452
FT /*tag= u
FT 4576..4718
FT /*tag= v
FT /label= 3' ITR
FT /note= "Inverted terminal repeat which is capable
FT of forming T-shaped hairpin structure"
FT
XX
PN WO200028061-A2.
XX
PD 18-MAY-2000.
XX
PF 02-NOV-1999; 99WO-US25694.
XX
PR 05-NOV-1998; 98US-0107114.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Wilson JM, Xiao W;
XX
DR WPI; 2000-376571/32.
DR P-PSDB; AAY71161, AAY71164, AAY71165, AAY71166, AAY71167, AAY71168,
DR AAY71169.
XX
PT Novel adeno-associated virus serotype 1 polynucleotide useful for
PT preparation of medicament for delivery of a transgene to a host -
XX Claim 1; Fig 1; 108pp; English.
PS
XX
CC The present sequence is an adeno-associated virus serotype 1 (AAV-1)
CC DNA characterised by two inverted terminal repeats (ITR) and open
CC reading frames for rep and capsid (cap) proteins. The rep reading frame
CC encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap
CC reading frame encodes three structural proteins, VP1, VP2 and VP3.
CC The AAV-1 sequence or its fragments particularly ITRs, rep and cap
CC coding regions, are useful in production of recombinant viral vectors
CC for gene delivery. These vectors can be used as gene therapy
CC vectors, vaccine vectors or antisense delivery vectors. The AAV-1 does
CC not induce the formation of neutralising antibodies specific to any
CC serotype of AAV hence is useful for transforming host cells, and in the
CC preparation of a medicament for the delivery of transgene to a host.
XX
SQ Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 other;

Alignment Scores:
Pred. No.: 5.03e-268 Length: 4718
Score: 2906.00 Matches: 534
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
```

```
DB: 21 Gaps: 0
US-09-807-802A-17 (1-534) x AAD00772 (1-4718)
QY 1 MetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20
Db 2829 ATGGCTTCAGGCGGTGGCGCACCATAATGGCAGACATAACGAAGCGCCGACGGAGTGGGT 2888
QY 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThr 40
Db 2889 AATGCCTCAGGAATTTGGCATTCGATTCCACATGGTGGGCGACAGAGTATCATCACCACC 2948
QY 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60
Db 2949 AGCACCCGCACCTGGGCTTGCCACCTACATAATACCACCTCTACAAGCAAAATCTCCAGT 3008
QY 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80
Db 3009 GCTTCAACGGGGCCAGCAACGACAACCACTACTTCGGCTACAGCACCCCTGGGGGTAT 3068
QY 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100
Db 3069 TTTGATTTCAACAGATTCCACTGCCACTTTTTCACCACGTGACTGGCAGCGACTCATCAAC 3128
QY 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120
Db 3129 AACAAATGGGGATTCCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAGTCAAG 3188
QY 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140
Db 3189 GAGGTCACGACGAATGATGGCGTCACAACCATCGCTAATAACCTTACCAGCACGGTTCAA 3248
QY 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160
Db 3249 GTCTTCTCGGACTCGGAGTACCAGCTTCCGTACGTCTCTGGCTCTGCGCACCAGGGCTGC 3308
QY 161 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180
Db 3309 CTCCTCCGTTCCCGCGGAGCTGTTCATGATTCGGCAATACGGCTACTGACGCTCAAC 3368
QY 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200
Db 3369 AATGGCAGCCCAAGCCGTGGAGCTTCACTCTTTTACTGCGTGAATATTTCCCTTCTCAG 3428
QY 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis 220
Db 3429 ATGCTGAGAACGGGCAACAACCTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 3488
QY 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240
Db 3489 AGCAGCTACGCGCACAGCCAGAGCCTGGACCGCTGATGAATCCTCTCATCGACCAATAC 3548
QY 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260
Db 3549 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCAAACAAAGGACTTGTCTG 3608
QY 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280
Db 3609 TTTAGCGTGGGTCTCCAGCTGGCATGTCTGTTTACGCCCCAAAACAACTGGCTACCTGGACCC 3668
QY 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300
Db 3669 TGTATCGGCAGCAGCGCGTCTTCTAAAACAAAACAGACAAACAAACAGCAATTTTACC 3728
QY 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320
Db 3729 TGGACTGTGCTTCAAAATATAACCTCAATGGCGTGAATCCATCATCAACCCCTGGCACT 3788
QY 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340
Db 3789 GCTATGGCCTCACACAAGACGACGAGACAAGTTCTTTTCCCATGAGCGGTGTCATGATT 3848
QY 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360
```

Db 3849 TTTGGAAAAGAGAGCGCGGAGCTTCAAACACTGCAATTTGGACAATGTCATGATTACAGAC 3908  
Qy 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380  
Db 3909 GAAGAGGAAATTAAAGCCACTAAACCTGTGGCCACCAGAAAGATTGGGACCGTGGCAGTC 3968  
Qy 381 AsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400  
Db 3969 AATTCCAGAGCAGCAGCAGACCCCTGGACCGGAGATGTCATGCTATGGGAGCATTA 4028  
Qy 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420  
Db 4029 CCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCAGGGTCCCATTTGGGCCAAATT 4088  
Qy 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440  
Db 4089 CCTCACAGATGGACACTTTACCCGTCTCCTCTTATGGCGGCTTTGGACTCAAGAAC 4148  
Qy 441 ProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460  
Db 4149 CCGCCTCCTCAGATCCTCATCAAAAACACGCGCTGTTCTCTGCGAATCCTCCGGCGGAGTTT 4208  
Qy 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480  
Db 4209 TCAGCTACAAAGTTTGTTTCATTTCATCACCCTATCTCCACAGGACAAGTGAGTGGAA 4268  
Qy 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 500  
Db 4269 ATTGAATGGGAGCTGCAGAAAGAAAACAGCAAGCGTGGAATCCCGAAGTGCAGTACACA 4328  
Qy 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520  
Db 4329 TCCAATATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAACAATGGACTTTTATACT 4388  
Qy 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
Db 4389 GAGCCTCGCCCATTTGGCACCCCGTTACCTTACCCGTCGCCCTG 4430

RESULT 5

AAF23749  
ID AAF23749 standard; DNA; 4683 BP.  
XX  
AC AAF23749;  
XX  
DT 28-MAR-2001 (first entry)  
XX  
DE AAV6 DNA sequence.  
XX  
KW AAV6; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;  
KW atherosclerosis; sickle cell anaemia; thalassaemia;  
KW blood clotting disorder; diabetes; ss.  
OS Adeno associated virus.  
XX  
PN US6156303-A.  
XX  
PD 05-DEC-2000.  
XX  
PF 11-JUN-1997; 97US-0873168.  
XX  
PR 11-JUN-1997; 97US-0873168.  
XX  
PA (UNIW ) UNIV WASHINGTON.  
XX  
PI Russell DW, Rutledge EA;  
XX  
DR WPI; 2001-060164/07.  
XX  
PT Adeno-associated virus serotype 6 and viral vector derived from it for  
PT gene therapy of cystic fibrosis, cancer, acquired immunodeficiency  
PT syndrome, sickle cell anemia, thalassemia and diabetes -  
XX

PS Claim 1; Fig 1; 50pp; English.  
XX  
CC The present invention relates to adeno-associated virus serotypes. The  
CC present sequence is the DNA sequence of one such serotype (AAV6). AAV6  
CC can be used to construct AAV viral vectors for use in gene therapy for a  
CC range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,  
CC sickle cell anaemia, thalassaemia, blood clotting disorders and diabetes.  
CC The AAV viral vectors have increased transduction efficiency of a  
CC particular host cell as the AAV virion containing the AAV vector genome  
CC can be modified to express a capsid protein of an AAV serotype that  
CC transduces the selected host cell.  
SQ Sequence 4683 BP; 1114 A; 1363 C; 1277 G; 929 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 6.41e-266 Length: 4683  
Score: 2884.00 Matches: 529  
Percent Similarity: 99.63% Conservative: 3  
Best Local Similarity: 99.06% Mismatches: 2  
Query Match: 99.24% Indels: 0  
DB: 22 Gaps: 0  
  
US-09-807-802A-17 (1-534) x AAF23749 (1-4683)  
  
Qy 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20  
Db 2814 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACATAACGAAGGCGCCGAGTGGGT 2873  
Qy 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40  
Db 2874 AATGCCTCAGGAAATTGGCATTTGGATTCCCATTTCCACATGGCTGGGCGACAGAGTCATCACCACC 2933  
Qy 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
Db 2934 AGCACCCGAACATGGGCTTGCCACCTATATAACAACCACTCTACAAGCAAAATCTCCAGT 2993  
Qy 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTyrGlyTyr 80  
Db 2994 GCTTCAACGGGGGCCAGCAACGACAACCACTACTTCGGCTACAGCACCCCTTGGGGGTAT 3053  
Qy 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100  
Db 3054 TTTGATTTCAACAGATTCCACTGCCATTTCTCACCACTGACTGGCAGCGACTCATCAAC 3113  
Qy 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120  
Db 3114 AACAAATTGGGATTCCGGCCCAAGAGACTCAACTTCAAGCTCTTCAACATCCCAAGTCAAG 3173  
Qy 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140  
Db 3174 GAGGTCACGACGAATGATGGCGTTCACGACCATCGCTAATAACCTTACCAGCACCGTTCAA 3233  
Qy 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160  
Db 3234 GTCTTCTCGGACTCGGAGTACCAGTTGCCGTACGTCTCGGCTCTGCGCACCCAGGCTGC 3293  
Qy 161 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
Db 3294 CTCCCTCCGTTCCCGCGGACGCTGTTTCATGATTCGCGAGTACGGCTACCTAACGCTCAAC 3353  
Qy 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
Db 3354 AATGGCAGCCAGGCAAGTGGACGGTTCATCTCTTACTGCCTGGAATATTTCCCATCGCAG 3413  
Qy 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis 220  
Db 3414 ATGCTGAGAACCGGCAATAAATTTACCTTTCAGCTACACCTTCGAGGACGTCGCTTTCCAC 3473  
Qy 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
Db 3474 AGCAGCTACGCGCACAGCCAGAGCCTGGACCGGCTGATGAATCTCTCATCGACCACTAC 3533  
Qy 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260

Db 3534 CTGTATTACCTGAACAGAACTCAGAATCAGTCCGGAAGTCCCAAAACAAGGACTTGCTG 3593  
Qy 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280  
Db 3594 TTTAGCCGGGGTCTCCAGCTGGCATGTCTGTTCCAGCCCAAAAAGTGGCTACCTGGACCC 3653  
Qy 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
Db 3654 TGTTACCGGCAGCAGCGCGTTTCTAAAACAAAACAGACAACAACAACAGCAACTTTACC 3713  
Qy 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleAsnProGlyThr 320  
Db 3714 TGGACTGGTGCTTCAAAATATAACCTTAATGGCGTGATCTATAATCAACCCCTGGCACT 3773  
Qy 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340  
Db 3774 GCTATGGCTCACACAAAGACGACAAGACAAGTCTTTCCCATGAGCGGTGTCATGATT 3833  
Qy 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360  
Db 3834 TTTGGAAAGGAGAGCGCGGAGCTTCAAAACACTGCATTGGACAATGTTCATGATCACAGAC 3893  
Qy 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380  
Db 3894 GAAGAGGAAATCAAAGCCACTAACCCCGTGGCCACCAGAAAGATTGGGACTGTGGCAGTC 3953  
Qy 381 AsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400  
Db 3954 AATCTCCAGACGACGACAGACCCTGGCAGCGGAGATGTGCTGTTATGGAGCCTTA 4013  
Qy 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420  
Db 4014 CCTGGAATGGTGGCAAGACAGACAGACGTATACCTGCAGGGTCTTATTTGGGCCAAAT 4073  
Qy 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440  
Db 4074 CCTCACACGGATGGACACTTTTACCCCGTCTCTCTCATGGCGGGCTTTGGACTTAAGCAC 4133  
Qy 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460  
Db 4134 CCGCCTCCTCAGATCCTCATCAAAAACACGCTGTTCTCGGAATCCTCCGGCAGAGTTT 4193  
Qy 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480  
Db 4194 TCGGCTACAAAGTTGCTTTCATTTCATCACCCAGTATTCACAGGACAAGTGAGCGTGGAG 4253  
Qy 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 500  
Db 4254 ATTGAATGGGAGTGCAGAAAGAAACAGCAACGCTGGAATCCCGAAGTGCAGTATACA 4313  
Qy 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520  
Db 4314 TCTAACTATGCAAAATCTGCCAACGTTGATTTCACTGTGGACAACATGGACTTTTACT 4373  
Qy 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
Db 4374 GAGCCTCGCCCCCATTTGGCACCCGTTTACCTCACCCGTCCTCCCTG 4415

RESULT 6  
AAF23748  
ID AAF23748 standard; DNA; 4722 BP.  
XX AAF23748;  
AC AAF23748;  
XX 28-MAR-2001 (first entry)  
DT 28-MAR-2001 (first entry)  
XX AAV3B DNA sequence.  
XX AAV3B; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;  
KW atherosclerosis; sickle cell anaemia; thalassaemia;  
KW blood clotting disorder; diabetes; ss.  
XX

OS Adeno associated virus.  
XX US6156303-A.  
PN 05-DEC-2000.  
PD 11-JUN-1997; 97US-0873168.  
XX 11-JUN-1997; 97US-0873168.  
PR (UNIW ) UNIV WASHINGTON.  
XX Russell DW, Rutledge EA;  
PI WPI; 2001-060164/07.  
XX Adeno-associated virus serotype 6 and viral vector derived from it for  
PT gene therapy of cystic fibrosis, cancer, acquired immunodeficiency  
PT syndrome, sickle cell anemia, thalassemia and diabetes -  
XX Example 2; Fig 1; 50pp; English.  
PS The present invention relates to adeno-associated virus serotypes. The  
XX present sequence is the DNA sequence of one such serotype (AAV3B). AAV3B  
CC can be used to construct AAV viral vectors for use in gene therapy for a  
CC range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,  
CC sickle cell anaemia, thalassaemia, blood clotting efficiency of a  
CC The AAV viral vectors have increased transduction efficiency of a  
CC particular host cell as the AAV virion containing the AAV vector genome  
CC can be modified to express a capsid protein of an AAV serotype that  
CC transduces the selected host cell.  
XX Sequence 4722 BP; 1204 A; 1297 C; 1243 G; 978 T; 0 other;  
SQ

Alignment Scores:  
Pred. No.: 5.86e-232 Length: 4722  
Score: 2530.00 Matches: 460  
Percent Similarity: 91.03% Conservative: 27  
Best Local Similarity: 85.98% Mismatches: 46  
Query Match: 87.06% Indels: 2  
DB: 22 Gaps: 2

US-09-807-802A-17 (1-534) x AAF23748 (1-4722)

Qy 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20  
Db 2814 ATGCTTCAGGCGGTGGCGCACCATAATGGCAGACAATAACGAGGGTGCCGATGGAGTGGT 2873  
Qy 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40  
Db 2874 AATCTCTCAGGAAATTTGGCATTTGGATTCCCAATGGTGGCGACAGAGTCATCACCACC 2933  
Qy 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
Db 2934 AGCACCAGAAACCTGGGCCCTGCCCACTTACAACAACCATCTCTACAAGCAAAATCTCCAGC 2993  
Qy 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80  
Db 2994 CAATCA---GGAGCTTCAAACGACAAACCACTACTTTTGGTACAGCACCCCTTTGGGGGTAT 3050  
Qy 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100  
Db 3051 TTTGACTTTAACAGATTCCACTGCCACTTCTCACCACGTCGACTGGCAGCGACTCATTAAC 3110  
Qy 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120  
Db 3111 AACAACTGGGATTCCCGGCCCAAGAACTCAGCTTCAAGCTTCTCAACATCCCAAGTTAA 3170  
Qy 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140  
Db 3171 GAGGTACGCAGAAACGATGGCAGCAGCAGACTATTGCCAATAACCTTACCAGCACGGTTCAA 3230  
Qy 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160



Db 3231 GTGTTACGGACTCGGAGTATCAGCTCCCGTACGTGCTCGGTCCGGCCACCAAGGCTGT 3290  
Qy 161 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
Db 3291 CTCCCGCGCTTCCAGCGGACGTCTTCATGGTCCCTCAGTATGGATACCTCACCTGAAC 3350  
Qy 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
Db 3351 AACGGAAGTCAAGCGGTGGACGCTCATCTTTTACTGCTGGAGTACTTCCCTTCGCAG 3410  
Qy 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPropheHis 220  
Db 3411 ATGCTAAGGACTGGAAATAACTTCCAATTCAGCTATACCTTCGAGGATGTACCTTTTCAC 3470  
Qy 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
Db 3471 AGCAGCTACGCTCACAGCCAGAGTTTGGATCGCTTGATGAATCCTCTTATTGATCAGTAT 3530  
Qy 241 LeuTyrTyrLeuAsnArgThrGln--AsnGlnSerGlySerAlaGlnAsnLysAspLeu 259  
Db 3531 CTGTACTACCTGAACAGAACGGAAGAACACCTCTGGAACAACCAACCAATCACGGCTG 3590  
Qy 260 LeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGly 279  
Db 3591 CTTTTTAGCCAGGCTGGGCTCAGTCTATGTCTTTGACGCCAGAAATTGGCTACCTGGG 3650  
Qy 280 ProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPhe 299  
Db 3651 CCCTGTACCGGCAACAGAGACTTTCAAAGACTGTCTAACGACAACAACAACAGTAACTTT 3710  
Qy 300 ThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGly 319  
Db 3711 CCTTGGACAGCGCCAGCAAAATATCATCTCAATGGCCGCGACTCGTGTGAATCCAGGA 3770  
Qy 320 ThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMet 339  
Db 3771 CCAGCTATGGCCAGTCAACAGGACGATGAAGAAAAATTTTCCCTATGCACGGCAATCTA 3830  
Qy 340 IlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThr 359  
Db 3831 ATATTGGCAAGAAGGACACCGCAAGTAACGCAGAAATAGATAATGTAATGATTACG 3890  
Qy 360 AspGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAla 379  
Db 3891 GATGAAGAAGAGATTCTGACCACCAATCCTGTGGCAACAGACAGTATGGAACCTGGCA 3950  
Qy 380 ValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAla 399  
Db 3951 AATAACTGCAGAGCTCAAATACAGTCCCACGACTAGAACTGTCAATGATCAGGGGCC 4010  
Qy 400 LeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLys 419  
Db 4011 TTACCTGGCATGGTGTGGCAAGATCGTGACGTGTACCTTCAAGGACCTATCTGGGCAAG 4070  
Qy 420 IleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLys 439  
Db 4071 ATTCTTCACACGGATGGACACTTTCATCCTTCTCCTCTGATGGAGGCTTTGGACTGAAA 4130  
Qy 440 AsnProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGlu 459  
Db 4131 CATCCGCTCCTCAAATCATGATGATCAAAAATACTCCGGTACCGGCAAAATCCTCCGACGACT 4190  
Qy 460 PheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerVal 479  
Db 4191 TTCAGCCCGGCCAAGTTTGTCTTATTATCACTCAGTACTCCACTGGACAGGTCAGCGTG 4250  
Qy 480 GluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyr 499  
Db 4251 GAAATTGAGTGGGAGCTACAGAAAGAAAACAGCAAAACGTTTGGAAATCCAGAGATTTCAGTAC 4310  
Qy 500 ThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyr 519

Db 4311 ACTTCCAACATAACAAGTCTGTTAATGTGGACTTTACTGTAGACACTAATGGTGTAT 4370  
Qy 520 ThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
Db 4371 AGTGAACCTCGCCCTATTGGAACCCCGGTATCTCACACGAAACTTG 4415  
RESULT 7  
AAD00834  
ID AAD00834 standard; DNA; 8178 BP.  
XX  
AC AAD00834;  
XX  
DT  
XX 08-SEP-2000 (first entry)  
DE Chimeric adeno-associated virus 2/3 capsid encoding helper plasmid.  
XX  
KW Adeno-associated virus; AAV2; AAV3; chimeric; gene therapy; plasmid;  
KW recombinant parvoviral vector; cellular tropism; rep gene; cap protein;  
KW capsid; gene delivery; VP1; VP2; VP3; ss.  
XX  
OS Chimeric - Adeno associated virus serotype 2.  
OS Chimeric - Adeno associated virus serotype 3.  
XX  
FH Key Location/Qualifiers  
FT CDS 251..2180  
FT /\*tag= a  
FT /product= "Rep 68"  
FT /note= "The specification states that this region  
FT encodes Rep 68 protein of AAV2. However, the stop  
FT position given in the specification for this CDS does  
FT not correspond to stop codon"  
FT 251..2114  
FT /\*tag= b  
FT /product= "Rep 78"  
FT /note= "The specification states that this region  
FT encodes Rep 78 protein of AAV2. However, the stop  
FT position given in the specification for this CDS does  
FT not correspond to stop codon"  
FT 923..2114  
FT /\*tag= c  
FT /product= "Rep 52"  
FT /note= "The specification states that this region  
FT encodes Rep 52 protein of AAV2. However, the stop  
FT position given in the specification for this CDS does  
FT not correspond to stop codon"  
FT 923..2180  
FT /\*tag= d  
FT /product= "Rep 40"  
FT /note= "The specification states that this region  
FT encodes Rep 40 protein of AAV2. However, the stop  
FT position given in the specification for this CDS does  
FT not correspond to stop codon"  
FT 2133..4342  
FT /\*tag= e  
FT /product= "VP1 cap protein"  
FT /note= "The specification states that this region  
FT encodes VP1 protein of AAV2. However, the stop  
FT position given in the specification for this CDS does  
FT not correspond to stop codon"  
FT 2544..4342  
FT /\*tag= f  
FT /product= "VP2 cap protein"  
FT /note= "The specification states that this region  
FT encodes VP2 protein of AAV2. However, the positions  
FT given in the specification for this CDS do not correspond  
FT to start or stop codons"  
FT 2739..4342  
FT /\*tag= g  
FT /product= "VP3 cap protein"  
FT /note= "The specification states that this region  
FT encodes VP3 protein of AAV2 in which loops 2-4 are  
FT replaced with the corresponding region from AAV3.  
FT However, the stop position given in the specification  
FT

FT misc\_feature for this CDS does not correspond to stop codon"  
FT 3184..4092 /\*tag= h  
FT /note= "AAV3 sequence"  
XX WO200028004-A1.  
PD 18-MAY-2000.  
XX 10-NOV-1999; 99WO-US26505.  
XX 10-NOV-1998; 98US-0107840.  
PR 10-MAR-1999; 99US-0123651.  
XX (UYNC-) UNIV NORTH CAROLINA.  
XX Rabinowitz JE, Samulski RJ, Xiao W;  
PI WPI; 2000-376523/32.  
XX Recombinant parvoviral vectors with altered packaging, tropisms and  
PT immunogenic properties, useful in gene therapy protocols -  
PT  
XX Example 27; Page 143-147; 153pp; English.  
XX The patent discloses modified parvovirus vectors with advantageous  
CC antigenic properties, packaging capabilities and cellular tropisms.  
CC These vectors can be used in standard recombinant DNA protocols e.g. gene  
CC therapy for delivering nucleic acids to cells. The present sequence is  
CC a helper plasmid encoding a chimeric adeno-associated virus serotype 2/3  
CC (AAV2/3) capsid. The sequence contains the AAV2 rep coding sequences,  
CC most of the AAV2 capsid coding sequences with the exceptions that  
CC loops 2-4 from the AAV2 VP3 subunit is replaced with the corresponding  
CC region from AAV3, in a pBluescript backbone. The rep  
CC coding sequence encodes four proteins responsible for replication,  
CC Rep 68, Rep 78, Rep 52 and Rep 40, and the capsid (cap) coding region  
CC encodes three structural proteins VP1, VP2 and VP3.  
CC Recombinant parvovirus comprising the chimeric capsid is useful  
CC for gene delivery.  
XX  
SQ Sequence 8178 BP; 2082 A; 2078 C; 2027 G; 1991 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 1.82e-230 Length: 8178  
Score: 2518.00 Matches: 458  
Percent Similarity: 90.65% Conservative: 27  
Best Local Similarity: 85.61% Mismatches: 48  
Query Match: 86.65% Indels: 2  
DB: 21 Gaps: 2  
  
US-09-807-802A-17 (1-534) x AAD00834 (1-8178)  
  
QY 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20  
DB 2739 ATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGGCGCCGACGGAGTGGGT 2798  
  
QY 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40  
DB 2799 AATTCCTCCGGAATTGGCATTGGCATTCACATGGATGGCGACAGAGTCATCACCACC 2858  
  
QY 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
DB 2859 AGCACCCGAACCTGGGCCCTGCCACCTACACACCACCCTCTACAAACAAATTTCCAGC 2918  
  
QY 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80  
DB 2919 CAATCA---GGAGCCTCGAACGACAATCACTATTGGCTACAGCACCCCTGGGGGTAT 2975  
  
QY 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100  
DB 2976 TTTGACTTCACAGATTCCACTGCCACTTTTCCACCAGTGGACTGGCAAAGACTCATCAAC 3035  
  
QY 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120

DB 3036 AACAACTGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAACATTCAAGTCAAA 3095  
  
QY 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140  
DB 3096 GAGGTCACGCAGAAATGACGGTACGACGACGAGATTGCCAATAAACCTTTACCAGACGGTTCAG 3155  
  
QY 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160  
DB 3156 GTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTGCTCGGGTCGGCGCACCAAGGCTGT 3215  
  
QY 161 LeuProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
DB 3216 CTCCCGCGGTTTCCAGCGGACGTCTTTCATGGTCCCTCAGTATGGATACCTCACCTGAAAC 3275  
  
QY 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
DB 3276 AACGGAAGTCAAGCGGTGGGACGCTCATCTTTTACTGCTGGAGTACTTCCCTTCGCAG 3335  
  
QY 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis 220  
DB 3336 ATGCTAAGGACTGGAAATAACTTCCAATTTCAGCTATACCTTCGAGGATGTACCTTTTCAC 3395  
  
QY 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
DB 3396 AGCAGCTACGCTCACAGCCAGAGTTTGGATCGCTTGATGAATCCTCTTATTGATCAGTAT 3455  
  
QY 241 LeuTyrTyrLeuAsnArgThrGln---AsnGlnSerGlySerAlaGlnAsnLysAspLeu 259  
DB 3456 CTGTACTACCTGAACAGACGCAAGGAACAACCTCTGGAACAACAACAATCACCGGTG 3515  
  
QY 260 LeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGly 279  
DB 3516 CTTTTTAGCCAGGCTGGGCTCAGTCTATGTCTTTGAGCCAGAGAAATTTGGTACCTACCTGGG 3575  
  
QY 280 ProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnSerAsnPhe 299  
DB 3576 CCCTGTACCGGCAACAGAGACTTTCAAAGACTGTCTAAGACACAACAACAGTAACCTT 3635  
  
QY 300 ThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGly 319  
DB 3636 CCTTGGACAGCGGCCAGCAATATCATCTCAATGGCCGCGACTCGCTGGTGAATCCAGGA 3695  
  
QY 320 ThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMet 339  
DB 3696 CCAGCTATGGCCAGTCAACAGACGATGAAGAAAAATTTTCCCTATGCACGGCAATCTA 3755  
  
QY 340 IlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThr 359  
DB 3756 ATATTTGGCAAGAGGGACAAACGGCAAGTAACGCGAGATTAGATAATGTATGATTACG 3815  
  
QY 360 AspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAla 379  
DB 3816 GATGAAGAAGAGATTTCGTACCAACCAATCCTGTGGCAACAGACGAGTATGGAACGTGGCA 3875  
  
QY 380 ValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAla 399  
DB 3876 AATAACTTGCAGAGCTCAAAATACAGCTCCCACGACTGGAACCTGTCATCATCAGGGGCC 3935  
  
QY 400 LeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLys 419  
DB 3936 TTACCTGGCATGGTGTGGCAAGATCGTGACGTGTACCTTCAAGACCTATCTGGGCAAG 3995  
  
QY 420 IleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLys 439  
DB 3996 ATTCCTCACACGGATGGACACTTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4055  
  
QY 440 AsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGlu 459  
DB 4056 CATCCGCTCCTCAAATCATGATCAAAAAATACTCCGGTACCTCGGAATCCTTCGACCACC 4115  
  
QY 460 PheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerVal 479

Db 4116 TTCAGTCGGCGAAAGTTTGCTTCCTCATCACAGTACTCCACGGGACAGGTCAGCGTG 4175

Qy 480 GluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyr 499  
|||||

Db 4176 GAGATCAGTGGGAGCTGCAGAGGAAACAGCAACGCTGGAATCCGAAATTCAGTAC 4235

Qy 500 ThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyr 519  
|||||

Db 4236 ACTTCCAACACTACAACAAGTCTGTTAATCGTGACTTACCCTGGATACATAATGGCGTGAT 4295

Qy 520 ThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
:::|||||

Db 4296 TCAGAGCCTCGCCCCATTGGCACCATACCTGACTCGTAATCTG 4340

RESULT 8

AAH26327

ID AAH26327 standard; DNA; 4072 BP.

XX AAH26327;

DT 02-OCT-2001 (first entry)

XX Adenovirus helper Ad cap2.

KW AAV; vector; adenovirus; helper virus; Ad cap2; loxP site;  
gene therapy; ds.

XX Chimeric - Mastadenovirus.

OS Chimeric - Adeno associated virus.

OS Chimeric - Human cytomegalovirus.

XX WO200155361-A2.

XX 02-AUG-2001.

PF 26-JAN-2001; 2001WO-US02709.

XX 26-JAN-2000; 2000US-0178536.

PA (CHIR ) CHIRON CORP.

XX Hardy SF;

DR WPI; 2001-483239/52.

XX Producing recombinant adeno-associated virus (rAAV) vector, by stably  
PT infecting eukaryotic host cell with rAAV vectors, helper adenovirus or  
PT herpes virus, recombinant adenovirus or herpes vectors -

PS Example 2; Page 58-59; 63pp; English.

XX The present sequence is that of adenovirus helper Ad cap2, made  
CC by Cre/lox recombination and expressing adeno-associated virus  
CC (AAV) VP1, 2,3 from the cytomegalovirus immediate early promoter  
CC of pAdlox, and containing a loxP site. Ad cap2 was used to  
CC demonstrate recombinant AAV (rAAV) production from virally  
CC transduced cells. The invention provides methods and compositions  
CC for producing rAAV vector particles by: (a) introducing into a host  
CC cell (i) AAV packaging plasmid pfloxAAV (see AAH26326), (ii) a  
CC recombinant viral vector encoding plasmid, and (iii) a plasmid  
CC encoding herpesvirus, cytomegalovirus or adenoviral functions, or  
CC a herpesvirus, cytomegalovirus or adenovirus itself, to produce  
CC flox AAV particles and rAAV particles; and (b) introducing into a  
CC second host cell (i) the rAAV particles or (a), (ii) a vector  
CC that directs expression of Cre, and (iii) a vector which directs  
CC expression of herpes virus, cytomegalovirus or adenovirus helper  
CC functions, such that rAAV vector particles are produced. The  
CC vectors are useful for in vivo or in vitro gene therapy and also  
CC for in vitro recombinant protein production.

XX Sequence 4072 BP; 1073 A; 1025 C; 992 G; 982 T; 0 other;

SQ Alignment Scores:

Pred. No.: 2.13e-227 Length: 4072  
Score: 2481.50 Matches: 445  
Percent Similarity: 90.26% Conservative: 37  
Best Local Similarity: 83.33% Mismatches: 51  
Query Match: 85.39% Indels: 1  
DB: 22 Gaps: 1

US-09-807-802A-17 (1-534) x AAH26327 (1-4072)

Qy 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20  
|||||

Db 2090 ATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGGCGCCGCGAGTGGGT 2149  
|||||

Qy 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40  
|||||

Db 2150 AATTCCTCGGAAATTGGCATTGGGATTCCACATGGATGGCGACAGAGTCATCACCACC 2209  
|||||

Qy 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
|||||

Db 2210 AGCACCCGAACTGGGCCCTGCCACCTACAAACACCACCTCTACAAACAAATTTCCAGC 2269  
|||||

Qy 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80  
|||||

Db 2270 CAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTGGGGGTAT 2326  
|||||

Qy 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100  
|||||

Db 2327 TTTGACTTCAACAGATTCCACTGCCACTTTTCCACACGTGACTGGCAAAGACTCATCAAC 2386  
|||||

Qy 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120  
|||||

Db 2387 AACAACTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAACATTCAAGTCAAA 2446  
|||||

Qy 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140  
|||||

Db 2447 GAGGTCACGCAGAAATGACGGTACGACGACGATTGCCAATAACCTTACCAGCACGGTTCAG 2506  
|||||

Qy 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160  
|||||

Db 2507 GTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTCTCGGCTCGGCGCATCAAGGATGC 2566  
|||||

Qy 161 LeuProPropheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
|||||

Db 2567 CTCCCGCGCTTCCCAGCAGACGCTCTTCATGGTGCCACAGTATGGATACCTCACCCCTGAAC 2626  
|||||

Qy 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
|||||

Db 2627 AACGGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCTGGAGTACTTTCCCTTCTCAG 2686  
|||||

Qy 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis 220  
|||||

Db 2687 ATGCTGCGTACCGGAAACAACCTTTACCTTCAGCTACACTTTTGAGGACGTTCTCTTCCAC 2746  
|||||

Qy 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
|||||

Db 2747 AGCAGTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGACCAGTAC 2806  
|||||

Qy 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260  
|||||

Db 2807 CTGTATTACTTGAGCAGAAACAAACACTCCAAGTGGAAACCAACGAGTCAAGGCTTCAG 2866  
|||||

Qy 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280  
|||||

Db 2867 TTTTCTCAGGCCGAGCGAGTGACATTCCGGACCACTAGGAAGTGGCTTCTCTGGACCC 2926  
|||||

Qy 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
|||||

Db 2927 TGTACCGCCAGCAGCGAGTATCAAGACATCTCGCGGATAACAACAACAGTGAATCTCG 2986  
|||||

Qy 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320  
|||||

Db 2987 TGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGCCCG 3046  
|||||





QY 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20  
Db 2809 ATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGCGCCGACGGAGTGGT 2868  
QY 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40  
Db 2869 AATTCTCGGAAATTGGCATTTGGCATTCGATTCACATGGATGGGCGACAGATCATCAACC 2928  
QY 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
Db 2929 AGCACCCGAACCTGGGCCCTGCCACCTACAACAACCACTCTACAACAATAATTTCCAGC 2988  
QY 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80  
Db 2989 CAATCA--GGAGCCTCGAACGACAATCACTACTTTTGGCTACAGCACCCCTTGGGGGTAT 3045  
QY 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100  
Db 3046 TTTGACTTCAACAGATTCCACTGGCACTTTTCAACACCGTACTGGCAAGACTCATCAAC 3105  
QY 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120  
Db 3106 AACAACTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAACAATTCAAGTCAA 3165  
QY 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140  
Db 3166 GAGGTACGCAGCAATGACGGTACGACGACGATTGGCCAATAACCTTACCAGCACGGTTCAG 3225  
QY 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160  
Db 3226 GIGTTTACTGACTCGGAGTACCGAGTACCACTCCCGTACGTCCTCGGCTCGGCGCATCAAGGATGC 3285  
QY 161 LeuProPropheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
Db 3286 CTCCCGCGGTTCCAGCAGACAGTCTTCATGGTGCCACAGTATGGATACCTCACCCCTGAAC 3345  
QY 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
Db 3346 AACGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCTGGAGTACTTTTCTCTCTCAG 3405  
QY 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis 220  
Db 3406 ATGCTGCTACCGGAAACAACATTACCTTACGCTACACATTTTGAGGACGTTTCTCTCCAC 3465  
QY 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
Db 3466 AGCAGTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGACCCAGTAC 3525  
QY 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260  
Db 3526 CTGTATTACTTGAGCAGAAACAACACTCCAAGTGGAAACCAACCGCAGTCAAGGCTTCAG 3585  
QY 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280  
Db 3586 TTTTCTCAGGCCGAGCGAGTGACATTCGGGACCACTAGGAACCTGGCTTCTCTGGACCC 3645  
QY 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
Db 3646 TGTTACCGCAGCAGCGAGTATCAAAAGACATCTGCGGATACAACAACAGTGAATACTCG 3705  
QY 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320  
Db 3706 TGGACTGGAGTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGCCCG 3765  
QY 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340  
Db 3766 GCCATGGCAAGCCACAAGGACGATGAAGAAAGTTTTTTCTCAGAGCGGGGTCTTCATC 3825  
QY 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360  
Db 3826 TTTGGGAAGCAAGGCTCAGAGAAACAATAATGTGACATTGAAGAGGTTCATGATTACAGAC 3885

QY 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380  
Db 3886 GAAAGAGGAAATCAGACAACCAATCCCGTGGCTACGGAGCAGTATGGTTCTGTATCTACC 3945  
QY 381 AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400  
Db 3946 AACCTCCAGAGAGGCAACAGACAGACAGTACCGCAGATGTCAACACACACAAGCGCTCTT 4005  
QY 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420  
Db 4006 CCAGGCATGGTCTCGCAGGACAGAGATGTGTACTTTCAGGGGCCCATCTCGGCAAGATT 4065  
QY 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440  
Db 4066 CCACACACGGACGACATTTTCACCCCTCTCCCTCATGGGTGGATTCCGACTTAAACAC 4125  
QY 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460  
Db 4126 CCTCCTCCACAGATTCTCATCAAGAACACCCCGTACCTCGGAATCCTTCGACCACCTTC 4185  
QY 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480  
Db 4186 AGTGGCGGCAAGTTTGTCTTCTTTCATCACACAGTACTCCACGGGACAGGTCAAGGTGGAG 4245  
QY 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 500  
Db 4246 ATCGAGTGGAGCTGCAGAAGGAAACAGCAACGCTGGAATCCCGAAATTCAGTACACT 4305  
QY 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520  
Db 4306 TCCAACCTACAACAGTCTGTTAATGTGGACTTTACTGTGGACACTAATGGCGTGTATTCA 4365  
QY 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
Db 4366 GAGCCTCGCCCCCATTTGGCACCAGATACCTGACTCGTAATCTG 4407

RESULT 10

AAF23750  
ID AAF23750 standard; DNA; 4679 BP.

AC AAF23750;

XX 28-MAR-2001 (first entry)

DE AAV2 DNA sequence.

XX AAV2; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;  
KW atherosclerosis; sickle cell anaemia; thalassaemia;  
KW blood clotting disorder; diabetes; ss.

OS Adeno associated virus.

XX US6156303-A.

XX 05-DEC-2000.

PF 11-JUN-1997; 97US-0873168.

PR 11-JUN-1997; 97US-0873168.

XX (UNIW ) UNIV WASHINGTON.

PI Russell DW, Rutledge EA;

XX WPI; 2001-060164/07.

XX Adeno-associated virus serotype 6 and viral vector derived from it for  
PT gene therapy of cystic fibrosis, cancer, acquired immunodeficiency  
PT syndrome, sickle cell anemia, thalassemia and diabetes -

XX Claim 7; Fig 1; 50pp; English.

XX The present invention relates to adeno-associated virus serotypes. The

CC present sequence is the DNA sequence of one such serotype (AAV2). AAV2  
CC can be used to construct AAV viral vectors for use in gene therapy for a  
CC range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,  
CC sickle cell anaemia, thalassaemia, blood clotting disorders and diabetes.  
CC The AAV viral vectors have increased transduction efficiency of a  
CC particular host cell as the AAV virion containing the AAV vector genome  
CC can be modified to express a capsid protein of an AAV serotype that  
CC transduces the selected host cell.  
XX  
SQ Sequence 4679 BP; 1197 A; 1263 C; 1255 G; 964 T; 0 other;

Alignment Scores:                    2.6e-227                    Length:                    4679  
Pred. No.:                    2481.50                    Matches:                    445  
Score:                    90.26%                    Conservative:                    37  
Percent Similarity:                    83.33%                    Mismatches:                    51  
Best Local Similarity:                    85.39%                    Indels:                    1  
Query Match:                    22                    Gaps:                    1  
DB:

US-09-807-802A-17 (1-534) x AAF23750 (1-4679)

|    |      |  |      |
|----|------|--|------|
| QY | 1    | MetAlaSerGlyGlyGlyAlaProMetAlaAspAsnGluGlyAlaAspGlyValGly      | 20   |
| Db | 2809 | ATGGCTACAGGAGTGGCGCACCAATGGCAGACAATAACGAGGGCGCGAGTGGGT         | 2868 |
| QY | 21   | AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThr      | 40   |
| Db | 2869 | AATTCTCCGGAATTTGGCATTTGGATTCCACATGGATGGCGACAGATCATCACCAC       | 2928 |
| QY | 41   | SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSer      | 60   |
| Db | 2929 | AGCACCCGAACCTGGGCCCTGCCACCCTACAAACACCCTCTACAAACAATTTCCAG       | 2988 |
| QY | 61   | AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr   | 80   |
| Db | 2989 | CAATCA---GGAGCCTCGAACGACAATCACTATTGGCTACAGCACCCCTTGGGGGTAT     | 3045 |
| QY | 81   | PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn   | 100  |
| Db | 3046 | TTTGACTTCAACAGATTCCACTGCCACTTTTACCACCGTACTGGCAAAAGACTCATCAAC   | 3105 |
| QY | 101  | AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys   | 120  |
| Db | 3106 | AACAACCTGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTAAACATTCAAGTCAAA  | 3165 |
| QY | 121  | GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln   | 140  |
| Db | 3166 | GAGGTACCGCAGAAATGACGGTACGACGAGATTGCCAATAACCTTACCAGCACGGTTCAG   | 3225 |
| QY | 141  | ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys   | 160  |
| Db | 3226 | GTGTTTACTGACTCGGAGTACCAGTCCCGTACGTCCTCGGCTCGGCGCATCAAGGATGC    | 3285 |
| QY | 161  | LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn   | 180  |
| Db | 3286 | CTCCCGCCGTTCCACGACAGCGTCTTTCATGGTGCCACAGTATGGATACCTCACCCCTGAAC | 3345 |
| QY | 181  | AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln   | 200  |
| Db | 3346 | AACGGGAGTCAGGCAGTAGGACGCTCTTCAATTTACTGCCTGGAGTACTTTCTCTTCTCAG  | 3405 |
| QY | 201  | MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis   | 220  |
| Db | 3406 | ATGCTGCGTACCGGAAACAACACTTTACCTTACGTACACTTTTGAGGACGTTCTCTTCCAC  | 3465 |
| QY | 221  | SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr   | 240  |
| Db | 3466 | AGCAGTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCTCTCTCATCGACCATGAC   | 3525 |
| QY | 241  | LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu   | 260  |
| Db | 3526 | CTGTATTACTTGACGAGACAACAACACTCCAAGTGGGAACCAACCCAGTCAGGCTTCAG    | 3585 |

|    |      |   |      |
|----|------|---|------|
| QY | 261  | PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro    | 280  |
| Db | 3586 | TTTTTCTCAGCCCGGAGCGAGTGACATTTCGGACCAGTCTAGGAACCTGGCTTCTCTGGACCC | 3645 |
| QY | 281  | CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr    | 300  |
| Db | 3646 | TGTTACCGCCAGCAGCGAGTATCAAAGACATCTGCGGATAACAACAACAGTGAATACTCG    | 3705 |
| QY | 301  | TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr    | 320  |
| Db | 3706 | TGGACTGGAGCTACCAAGTACCACCTCAATGGCAGACACTCTCTGGTGAATCCGGGCCCCG   | 3765 |
| QY | 321  | AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle    | 340  |
| Db | 3766 | GCCATGGCAAGCCACACAGGACGATGAAGAAAAGTTTTCCTCAGAGCGGGTTCATC        | 3825 |
| QY | 341  | PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp    | 360  |
| Db | 3826 | TTTGGGAAGCAAGGCTCAGAGAAAACAAATGTGGACATTTGAAAAGGTTCATGATTACAGAC  | 3885 |
| QY | 361  | GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal    | 380  |
| Db | 3886 | GAAAGGAAATCAGGACAAACCAATCCCGTGGCTACGAGCAGTATGGTTCTGTATCTACC     | 3945 |
| QY | 381  | AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu    | 400  |
| Db | 3946 | AACCTCCAGAGAGGCAACAGACAGCTACCGCAGATGTCAACACACACAAAGCGTCTT       | 4005 |
| QY | 401  | ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaValIle    | 420  |
| Db | 4006 | CCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGCAAAAGATT   | 4065 |
| QY | 421  | ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn    | 440  |
| Db | 4066 | CCACACACGACGCGACATTTTCACCCCTCTCCCTCATGGGTGGATTTCGGACTTAAACAC    | 4125 |
| QY | 441  | ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe    | 460  |
| Db | 4126 | CCTCCTCCACAGATTCTCATCAAGAAACACCCCGGTACTCTGCGAATCCTTCGACCACCTTC  | 4185 |
| QY | 461  | SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu    | 480  |
| Db | 4186 | AGTCCGGCAAAAGTTTGCTTCTTCATCACACAGTACTCCACGGGACAGGTCAGCGTGGAG    | 4245 |
| QY | 481  | IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr    | 500  |
| Db | 4246 | ATCGAGTGGGAGCTGCAGAGGAAACAGCAACGCTGGAAATCCCGAAATTCAGTACACT      | 4305 |
| QY | 501  | SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr    | 520  |
| Db | 4306 | TCCAACTACAACAAGTCTGTAAATGTGGACTTTTACTGTGGACACTAATGGCGTGTATTCA   | 4365 |
| QY | 521  | GluProArgProIleGlyThrArgTyrLeuThrArgProLeu                      | 534  |
| Db | 4366 | GAGCCTCGCCCCCATTTGGCACCAAGATACCTGACTCGTAAATCTG                  | 4407 |

RESULT 11

|          |  |
|----------|--|
| ABK89694 |  |
| ID       | ABK89694 standard; DNA; 4679 BP.                                     |
| XX       |  |
| AC       | ABK89694;  |
| XX       |  |
| DT       | 05-NOV-2002 (first entry)  |
| XX       |  |
| DE       | Adeno-associated virus 2 (AAV2) vector.                              |
| XX       |  |
| KW       | Adeno-associated virus 2 vector; AAV2; ds; cyclic; circular; cancer; |
| KW       | VPI capsid; heparin-sulphate proteoglycan; vaccine; immune response; |
| XX       | ovarian cancer.  |
| OS       | Adeno-associated virus 2.  |
| XX       |  |
| FH       | Key                    Location/Qualifiers                           |





|||||  
4066 CCACACACGACGACATTTTACCCCTCTCCCTCATGGTGGATTGGACTTAAACAC 4125  
441 ProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460  
4126 CCTCTCCACAGATTCTCATCAAGAACACCCCGGTACCTCGGAATCCTTCGACCACCTTC 4185  
461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480  
4186 AGTGGGCAAGTTTGCTTCTTCATCACACAGTACTCCACGGACAGGTCAGCGTGGAG 4245  
481 IleGluTyrGluLeuGlnLysGluAsnSerLysArgTyrAsnProGluValGlnTyrThr 500  
4246 ATCGAGTGGAGCTGCAGAGGAAACAGCAACGCTGGAATCCCGAAATTCAGTACACT 4305  
501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520  
4306 TCCAACTACAACAAGTCTGTTAATGTGGACTTTACTGTGGACACTAATGGCGTGTATTCA 4365  
521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
4366 GAGCCTCGCCCATGGCCACCAGATACCTGACTCGTAATCTG 4407

RESULT 12  
ABV76133  
ID ABV76133 standard; DNA; 4679 BP.  
XX  
AC ABV76133;  
XX  
XX  
DT 07-MAR-2003 (first entry)  
XX

Adeno associated virus type 2 coat protein nucleic acid.  
Gene therapy; vector; hepatitis B virus; cardiovascular disease;  
heart; cardiant; vasotropic; antiarrhythmic; antiarteriosclerotic;  
coat protein; gene; ss.

Adeno associated virus type 2.  
WO200287594-A1.  
XX  
PN  
XX  
PD 07-NOV-2002.  
XX  
PF 30-APR-2002; 2002WO-US13644.  
XX  
PR 30-APR-2001; 2001US-287423P.  
XX

(REGC ) UNIV CALIFORNIA.  
Chien KR, Hoshijima M;  
XX  
WPI; 2003-111844/10.  
XX  
Novel non-viral vector comprises vesicular membrane with hepatitis B  
envelope protein with cardiac targeting sequence, and nucleotide  
sequence for gene therapy useful for treating, e.g., heart failure,  
arrhythmia and atherosclerosis -  
Disclosure; Page 24-27; 53pp; English.  
XX

The present sequence is that of an adeno-associated virus (AAV) type  
2 coat protein nucleic acid. The invention provides a non-viral  
vesicle vector for the delivery of nucleic acid to various cardiac  
cell types. The vesicle vector contains the hepatitis B virus  
envelope protein in which at least part of the liver targeting  
sequence is deleted and replaced with a specific cardiac cell  
targeting sequence. For example, the loop IV region of the AAV  
binds to heparin sulfate proteoglycans on the surface of  
cardiomyocytes, and can be used to target the vector to  
cardiomyocytes. The vesicle vector can be delivered intravenously  
or intra-arterially rather than by more invasive methods such as  
direct cardiac injection. It can be used to deliver gene products  
to replace or enhance expression of proteins for treatment of heart

CC failure, arrhythmia, reperfusion injury, atherosclerosis, to  
CC promote angiogenesis, etc. The vesicles are highly stable and can  
CC be produced in large quantities, making them ideal for gene therapy.  
XX  
SQ Sequence 4679 BP; 1198 A; 1262 C; 1255 G; 964 T; 0 other;

Alignment Scores:  
Pred. No.: 2.6e-227 Length: 4679  
Score: 2481.50 Matches: 445  
Percent Similarity: 90.26% Conservative: 37  
Best Local Similarity: 83.33% Mismatches: 51  
Query Match: 85.39% Indels: 1  
DB: 25 Gaps: 1

US-09-807-802A-17 (1-534) x ABV76133 (1-4679)

Qy 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20  
Db 2809 ATGGCTACAGGCAGTGGCGCACCAATGGCAGACATAACGAGGGCGCGGAGTGGGT 2868  
Qy 21 AsnAlaSerGlyAsnTyrHisCysAspSerThrTyrLeuGlyAspArgValIleThrThr 40  
Db 2869 AATTCTCGGAAATTGGCATTCGATTCCCATTCACATGGATGGCGGACAGAGTCAACACC 2928  
Qy 41 SerThrArgThrTyrAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
Db 2929 AGCACCCGAACCTGGGCCCTGCCACCTACACCAACACACCTCTACAAACAAATTTCCAGC 2988  
Qy 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTyrGlyTyr 80  
Db 2989 CAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTTGGGGGTAT 3045  
Qy 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTyrGlnArgLeuIleAsn 100  
Db 3046 TTTCGACTTCAACAGATTCCACTGCCACTTTTTCACCACCGTACTGGCAAAGACTCATCAAC 3105  
Qy 101 AsnAsnTyrGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120  
Db 3106 AACAACTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAACATTCAAGTCAA 3165  
Qy 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140  
Db 3166 GAGGTACGCAGAAATGACGGTACGACGACGATTGCCAATAACCTTACCAGCACGGTTTCAG 3225  
Qy 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160  
Db 3226 GTGTTTACTGACTCGGAGTACAGCTCCCGTACGTCTCGGCTCGGCGCATCAAGGATGC 3285  
Qy 161 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
Db 3286 CTCCCGCGTTCCACGACAGACGCTTTCATGGTGCCACAGTATGATGATACCTCACCCCTGAAC 3345  
Qy 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
Db 3346 AACGGGAGTCAGGCAGTAGGACGCTTTCATTTTACTGCTGGAGTACTTTCTTCTCTCAG 3405  
Qy 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPropheHis 220  
Db 3406 ATGCTGCGTACCGGAAACAACACTTTACCTTCAGCTACACTTTTGAGGACGTTCTTTCCAC 3465  
Qy 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
Db 3466 AGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGATGATCTCTCATCGACCAGTAC 3525  
Qy 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260  
Db 3526 CTGTATTACTTGAGCAGAACAAACACTCCAAGTGGAAACCAACCGCAGTCAAGGCTTCAG 3585  
Qy 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTyrLeuProGlyPro 280  
Db 3586 TTTTCTCAGGCGGAGCGAGTGACATTTCGGGACCAGTCTAGGAACCTGGCTTCTTGGACCC 3645  
Qy 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300

Db 3646 ||||| TGTTACCGCAGCAGCGAGTATCAAGACATCTCGGATAACAACAACAGTGAATACTCG 3705

Qy 301 TtpThrGlyAlaSerLysTyrAsnLeuAenGlyArgGluSerIleAenProGlyThr 320

Db 3706 TGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGCCG 3765

Qy 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetile 340

Db 3766 GCCATGGCAAGCCACAAGGACGATGAAGAAAGTTTTTCCTCAGAGCGGGTTCTCATC 3825

Qy 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAenValMetIleThrAsp 360

Db 3826 TTTGGGAAGCAAGGCTCAGAGAAACAAATGTGGACATTGAAGAAGTCAATGATACAGAC 3885

Qy 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380

Db 3886 GAAGAGGAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTATGGTTCTGTATCTACC 3945

Qy 381 AsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400

Db 3946 AACCTCCAGAGAGGCAACAGACACAGCAGCTACCGCAGATGTCAACACACAAAGCGTTCTT 4005

Qy 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTipAlaLysIle 420

Db 4006 CCAGGCATGCTTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGCAAGATT 4065

Qy 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAen 440

Db 4066 CCACACACGGACGGACATTTTTCACCCCTCTCCCTCATGGTGGATTGGACTTAAACAC 4125

Qy 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460

Db 4126 CCTCCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAATCCTTCGACCACCTTC 4185

Qy 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480

Db 4186 AGTCGGGCAAGTTTGCTTCTTCATCACACAGTACTCCACGGACAGGTCAGCGTGGAG 4245

Qy 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 500

Db 4246 ATCGAGTGGAGCTGCAGAGGAAACAGCAACCGCTGGAATCCGGAATTCAGTACACT 4305

Qy 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAenAsnGlyLeuTyrThr 520

Db 4306 TCCAACTACAACAAGTCTGTTAATGTGACTTTACTGTGGACACTAATGGCGTGTATTCA 4365

Qy 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534

Db 4366 GAGCCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTG 4407

RESULT 13

AAH26326

ID AAH26326 standard; DNA; 7557 BP.

XX

AC AAH26326;

XX

XX 02-OCT-2001 (first entry)

DT

XX Adeno-associated virus packaging plasmid pfloxAAV.

DE

XX AAV; vector; pfloxAAV2; packaging cell line; gene therapy; ds.

KW

XX Adeno associated virus.

OS

XX WO200155361-A2.

PN

XX 02-AUG-2001.

PD

XX 26-JAN-2001; 2001WO-US02709.

PF

XX 26-JAN-2000; 2000US-0178536.

PR

XX

PA (CHIR ) CHIRON CORP.

XX

PI Hardy SF;

XX

DR WPI; 2001-483239/52.

XX

PT Producing recombinant adeno-associated virus (rAAV) vector, by stably

PT infecting eukaryotic host cell with rAAV vectors, helper adenovirus or

PT herpes virus, recombinant adenovirus or herpes vectors -

XX

PS Example 1; Page 54-57; 63pp; English.

XX

CC The present sequence is that of adeno-associated virus (AAV)

CC packaging plasmid pfloxAAV2. The following AAV packaging plasmids

CC were prepared: PKSrepCap, a non-replicating control plasmid

CC containing rep and cap genes but no inverted terminal repeats

CC (ITRs); pAV2 (see AAH26325), containing a wild-type AAV genome; and

CC pfloxAAV, a version of pAV2 with 2 loxP sites inserted such they

CC flank the rep and cap genes and thus separate the ITRs from these

CC genes. Each was combined with pCMV GFP (see AAH26324), a plasmid

CC with a green fluorescent protein expressing recombinant AAV (rAAV)

CC genome, and used to transfect 293 or Cre8 cells. After 6 hr, the

CC medium was changed and wild-type adenovirus type 5 was added.

CC Virus particles were harvested 3 days later. Plasmid pAV2

CC produced mostly AAV and a low yield of GFP vector with no effect

CC of Cre recombinase. pfloxAAV packaged as much GFP vector as the

CC control plasmid and while the amount of GFP vector was not

CC affected by Cre recombinase, the amount of floxAAV in the product

CC was reduced to 1/10 of the 293 value by the action of Cre

CC recombinase. In 293 cells it was evident that the GFP vector did

CC not compete with wild-type AAV for replication or packaging, while

CC the GFP vector was able to complete against floxAAV. Methods

CC and compositions are provided for producing rAAV vector particles

CC by: (a) introducing into a host cell (i) pfloxAAV, (ii) a

CC recombinant viral vector encoding plasmid, and (iii) a plasmid

CC encoding herpesvirus, cytomegalovirus or adenoviral functions, or

CC a herpesvirus, cytomegalovirus or adenovirus itself, to produce

CC flox AAV particles and rAAV particles; and (b) introducing into a

CC second host cell (i) the rAAV particles or (a), (ii) a vector

CC that directs expression of Cre, and (ii) a vector which directs

CC expression of herpes virus, cytomegalovirus or adenovirus helper

CC functions, such that rAAV vector particles are produced. The

CC vectors are useful for in vivo or in vitro gene therapy and also

CC for in vitro recombinant protein production.

XX

SQ Sequence 7557 BP; 1970 A; 1940 C; 1953 G; 1694 T; 0 other;

Alignment Scores:

Pred. No.: 5.16e-227 Length: 7557

Score: 2481.50 Matches: 445

Percent Similarity: 90.26% Conservative: 37

Best Local Similarity: 83.33% Mismatches: 51

Query Match: 85.39% Indels: 1

DB: 22 Gaps: 1

US-09-807-802A-17 (1-534) x AAH26326 (1-7557)

Qy 1 MetAlaSerGlyGlyAlaProMetAlaAspAenGluGlyAlaAspGlyValGly 20

Db 2837 ATGGCTACAGGCAGTGGCGCACCAATGGCAGACATAACGAGGGCGCGAGTGGGT 2896

Qy 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40

Db 2897 AATTCTCGGAAATTTGGCATTTCCGATTCCACATGGATGGCGCAGAGTCATCACACC 2956

Qy 41 SerThrArgThrTipAlaLeuProThrTyrAsnAenHisLeuTyrLysGlnIleSerSer 60

Db 2957 AGCACCCGAACCTGGGCCCTGCCACCTACACACACCACCTCTACAAACAATTTCCAGC 3016

Qy 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTipGlyTyr 80

Db 3017 CAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTGGGGGTAT 3073





CC and compositions are provided for producing rAAV vector particles  
CC by: (a) introducing into a host cell (i) pflxoAAV, (ii) a  
CC recombinant viral vector encoding plasmid, and (iii) a plasmid  
CC encoding herpesvirus, cytomegalovirus or adenoviral functions, or  
CC a herpesvirus, cytomegalovirus or adenovirus itself, to produce  
CC flox AAV particles and rAAV particles; and (b) introducing into a  
CC second host cell (i) the rAAV particles or (a), (ii) a vector  
CC that directs expression of Cre, and (iii) a vector which directs  
CC expression of herpes virus, cytomegalovirus or adenovirus helper  
CC functions, such that rAAV vector particles are produced. The  
CC vectors are useful for in vivo or in vitro gene therapy and also  
CC for in vitro recombinant protein production.  
XX  
SQ Sequence 8698 BP; 2117 A; 2368 C; 2306 G; 1907 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 6,31e-227 Length: 8698  
Score: 2481.50 Matches: 445  
Percent Similarity: 90.26% Conservative: 37  
Best Local Similarity: 83.33% Mismatches: 51  
Query Match: 85.39% Indels: 1  
DB: 22 Gaps: 1  
  
US-09-807-802A-17 (1-534) x AAH26325 (1-8698)  
  
Qy 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValcily 20  
Db 2809 ATGGCTACAGGAGTGGCGCACCAATGGCAGACAATAACGAGGGCGCCGAGTGGGT 2868  
  
Qy 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40  
Db 2869 AATTCTTCGGGAATTTGGCATTGGCATTCACATGGATGGGCGCAGAGTCATCACCAC 2928  
  
Qy 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
Db 2929 AGCACCGAACCTGGGCCCTGCCACCTACAAACAACCACTTACAAACAATTTCCAGC 2988  
  
Qy 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80  
Db 2989 CAATCA---GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTGGGGGTAT 3045  
  
Qy 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100  
Db 3046 TTTGACTTCAACAGATTCCACTGCCACTTTTCCACCAGTACTGCGCAAGACTCATCAAC 3105  
  
Qy 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnVallys 120  
Db 3106 AACAACTGGGATTCCGACCCCAAGAGACTCACTTCAAGCTCTTTAATCAAGTCAAA 3165  
  
Qy 121 GluValThrThrAsnAspGlyValThrIleAlaAsnAsnLeuThrSerThrValGln 140  
Db 3166 GAGGTACGCAAGATGACGGTACGACGACGATTGCCAATAACCTTACCAGCACGGTTTAC 3225  
  
Qy 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160  
Db 3226 GTGTTTACTGACTCGGAGTACCAGTCCCGTACGTCCTCGGCTCGGCGCATCAAGGATGC 3285  
  
Qy 161 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
Db 3286 CTCCCGCGGTTCGCCAGCAGCGTCTTCATGGTGCCACAGTATGGGATACCTCACCCCTGAAC 3345  
  
Qy 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
Db 3346 AACGGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCTGGAGTACTTCTCTTCTCAG 3405  
  
Qy 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPropheHis 220  
Db 3406 ATGCTGCGTACCGGAACAACCTTTACCTTCAGCTACACTTTTGGAGGACGTTCTCTTCCAC 3465  
  
Qy 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
Db 3466 AGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGACCGAGTAC 3525

Qy 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260  
Db 3526 CTGTATTACTTGAGCAGAAACAACACTCCAAGTGGGAACCAACCAGCAGTCAAGGCTTCAG 3585  
  
Qy 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280  
Db 3586 TTTTCTCAGGCCGAGCGAGTGACATTCCGGGACGAGTCTAGGAACCTGGCTTCTCTGGACCC 3645  
  
Qy 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
Db 3646 TGTACCAGCCAGCAGCGAGTATCAAGACATCTGCGGATAACAACAACAGTGAATACTCG 3705  
  
Qy 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleAsnProGlyThr 320  
Db 3706 TGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGCCCG 3765  
  
Qy 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340  
Db 3766 GCCATGGCAAGCCACAAGGACGATGAAGAAAAGTTTTCCTCAGAGCGGGTTCATC 3825  
  
Qy 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360  
Db 3826 TTTGGGAAGCAAGGCTCAGAGAAAACAATGTGACATTGAAAGGTTCATGATTACAGAC 3885  
  
Qy 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380  
Db 3886 GAAGAGGAATCAGGACAACCAATCCCCTGGCTACCGAGCAGTATGGTTCTGTATCTACC 3945  
  
Qy 381 AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400  
Db 3946 AACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAACACACAAGCGTCTT 4005  
  
Qy 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaValIle 420  
Db 4006 CCAGGCATGGTCTGCGAGGACAGAGATGTGTACCTTCAGGGGCCCATCTCGGCAAGATT 4065  
  
Qy 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440  
Db 4066 CCACACACGGACGGACATTTTCACCCCTCTCCCTCATGGTGGATTCGGACTTAAACAC 4125  
  
Qy 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460  
Db 4126 CCTCTCCACAGATTCTCATCAAGAACAACCCCGGTACCTCGGAATCCTTCGCCACCTTC 4185  
  
Qy 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480  
Db 4186 AGTGGGCAAGTTTGCTTCTTCTCATCACAGTACTCCACGGGACAGGTCAAGCTGGAG 4245  
  
Qy 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 500  
Db 4246 ATCGAGTGGGAGCTGCAGAAGGAAACAGCAACCGCTGGGAATCCCGAAATTCAGTACACT 4305  
  
Qy 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520  
Db 4306 TCCAACACTACAACAAGTCTGTTAATGTGGACTTTTACTGTGGACACTAATGGCGTGTATTCA 4365  
  
Qy 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
Db 4366 GAGCCTCGCCCCCATTTGGCACCAGATACCTGACTCGTAATCTG 4407

RESULT 15

AAT09008  
ID AAT09008 standard; DNA; 4680 BP.

XX  
AC AAT09008;

XX  
DT 05-JUL-1996 (first entry)

XX  
DE Wild-type adeno-associated virus 2 genome.

XX  
KW adeno-associated virus 2; AAV2; inverted terminal repeat; viral vector;  
KW DNA delivery; immunodeficiency virus protein; immunity; human; simian;  
KW neurological disorder; therapy; ss.





|    |      |   |      |
|----|------|---|------|
| Qy | 461  | SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGln-ValSerValG1  | 480  |
| Db | 4186 | AGTGGGCAAAAGTTTGCTTCCTTCATCACACAGTACTCCACGGACACGGTCAGCGTGA    | 4245 |
| Qy | 480  | uileGluTrpGluLeuGlnLysGluAsnSerLysArgTrrPaenProGluValGlnTyrTh | 500  |
| Db | 4246 | GATCGAGTGGGAGCTGCAGAGGAAAAACAGCAACGCTGGATCCCGAAATTCAGTACAC    | 4305 |
| Qy | 500  | rSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrTh  | 520  |
| Db | 4306 | TTCCCACTACAACAAGTCTGTTAATCGTGGACTTACCGTGGATACTAATGGCGTGATTC   | 4365 |
| Qy | 520  | rGluProArgProIleGlyThrArgTyrLeuThrArgProLeu                   | 534  |
| Db | 4366 | AGAGCCTCGCCCCCATTTGGCACCAGATACCTGACTCGTAATCTG                 | 4408 |

Search completed: February 14, 2004, 17:23:51  
Job time : 475.286 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.  
  
OM protein - nucleic search, using frame\_plus\_p2n model  
  
Run on: February 14, 2004, 16:11:58 ; Search time 78.8571 Seconds  
(without alignments)  
2988.931 Million cell updates/sec  
  
Title: US-09-807-802A-17  
Perfect score: 2906  
Sequence: 1 MASGGGAPMADNNEGADGVG.....NNGLYTEPRPIGTRYLTRPL 534  
  
Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0  
  
Searched: 569978 seqs, 220691566 residues 1139956  
Total number of hits satisfying chosen parameters:  
  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09807802/runat\_11022004\_175609\_15955/app\_query.fasta\_1.2389  
-DB=Issued Patents\_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USRF=US09807802 @CGN 1 1 131 @runat\_11022004\_175609\_15955 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*  
  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| SUMMARIES  |        |             |                   |
|------------|--------|-------------|-------------------|
| Result No. | Score  | Query Match | Description       |
| 1          | 2518   | 86.6        | Sequence 5, Appli |
| 2          | 2481.5 | 85.4        | Sequence 4, Appli |
| 3          | 2481.5 | 85.4        | Sequence 3, Appli |
| 4          | 2481.5 | 85.4        | Sequence 2, Appli |
| 5          | 2439.5 | 83.9        | Sequence 1, Appli |
| 6          | 2439.5 | 83.9        | Sequence 1, Appli |
| 7          | 2439.5 | 83.9        | Sequence 1, Appli |
| 8          | 2439.5 | 83.9        | Sequence 1, Appli |
| 9          | 2258.5 | 77.7        | Sequence 2, Appli |
| 10         | 1690.5 | 58.2        | Sequence 19, Appl |
| 11         | 1690.5 | 58.2        | Sequence 17, Appl |
| 12         | 1690.5 | 58.2        | Sequence 5, Appli |

|    |        |      |      |   |                    |                    |
|----|--------|------|------|---|--------------------|--------------------|
| 13 | 1690.5 | 58.2 | 4767 | 4 | US-09-532-594B-1   | Sequence 1, Appli  |
| 14 | 1690.5 | 58.2 | 7214 | 4 | US-09-438-268-1    | Sequence 1, Appli  |
| 15 | 491.5  | 16.9 | 2271 | 4 | US-09-438-268-3    | Sequence 3, Appli  |
| 16 | 238    | 8.2  | 1740 | 1 | US-07-969-213-1    | Sequence 1, Appli  |
| 17 | 212    | 7.3  | 1820 | 6 | 5223424-12         | Patent No. 5223424 |
| 18 | 209.5  | 7.2  | 2062 | 6 | 5223424-3          | Patent No. 5223424 |
| 19 | 194    | 6.7  | 2254 | 3 | US-08-552-369-1    | Sequence 1, Appli  |
| 20 | 189    | 6.5  | 5049 | 1 | US-08-336-345-1    | Sequence 1, Appli  |
| 21 | 189    | 6.5  | 5049 | 1 | US-08-336-345-2    | Sequence 2, Appli  |
| 22 | 189    | 6.5  | 5049 | 2 | US-08-647-655-1    | Sequence 1, Appli  |
| 23 | 189    | 6.5  | 5049 | 2 | US-08-647-655-2    | Sequence 2, Appli  |
| 24 | 183    | 6.3  | 1752 | 3 | US-09-022-949-1    | Sequence 1, Appli  |
| 25 | 178    | 6.1  | 1755 | 2 | US-08-317-785-1    | Sequence 1, Appli  |
| 26 | 119    | 4.1  | 3288 | 4 | US-09-107-532A-201 | Sequence 201, App  |
| 27 | 112.5  | 3.9  | 2394 | 3 | US-09-319-989-7    | Sequence 7, Appli  |
| 28 | 111    | 3.8  | 2000 | 1 | US-08-469-202-25   | Sequence 25, Appl  |
| 29 | 111    | 3.8  | 2000 | 2 | US-08-484-434C-32  | Sequence 32, Appl  |
| 30 | 111    | 3.8  | 2000 | 4 | US-09-384-361-32   | Sequence 32, Appl  |
| 31 | 109    | 3.8  | 3561 | 1 | US-08-485-568A-3   | Sequence 3, Appli  |
| 32 | 109    | 3.8  | 3561 | 1 | US-08-357-698-5    | Sequence 5, Appli  |
| 33 | 109    | 3.8  | 3561 | 2 | US-08-590-554A-3   | Sequence 3, Appli  |
| 34 | 109    | 3.8  | 3561 | 2 | US-09-184-223-3    | Sequence 3, Appli  |
| 35 | 109    | 3.8  | 3561 | 5 | PCT-US93-12682-5   | Sequence 5, Appli  |
| 36 | 108    | 3.7  | 1988 | 1 | US-08-469-202-26   | Sequence 26, Appl  |
| 37 | 108    | 3.7  | 1988 | 2 | US-08-484-434C-33  | Sequence 33, Appl  |
| 38 | 108    | 3.7  | 1988 | 4 | US-08-384-361-33   | Sequence 33, Appl  |
| 39 | 106    | 3.6  | 3155 | 2 | US-08-231-193A-13  | Sequence 13, Appl  |
| 40 | 106    | 3.6  | 3155 | 2 | US-08-486-273A-13  | Sequence 13, Appl  |
| 41 | 106    | 3.6  | 3155 | 3 | US-08-480-474-13   | Sequence 13, Appl  |
| 42 | 106    | 3.6  | 3155 | 3 | US-08-940-086A-13  | Sequence 13, Appl  |
| 43 | 106    | 3.6  | 3155 | 4 | US-08-940-035A-13  | Sequence 13, Appl  |
| 44 | 106    | 3.6  | 3155 | 4 | US-08-935-105A-13  | Sequence 13, Appl  |
| 45 | 106    | 3.6  | 3155 | 4 | US-09-648-797-13   | Sequence 13, Appl  |

ALIGNMENTS

RESULT 1  
US-09-438-268-5  
; Sequence 5, Application US/09438268  
; Patent No. 6491907  
; GENERAL INFORMATION:  
; APPLICANT: Rabinowitz, Joseph E.  
; APPLICANT: Samulski, Richard J  
; APPLICANT: Xiao, Weidong  
; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING  
; TITLE OF INVENTION: THE SAME  
; FILE REFERENCE: 5470-186  
; CURRENT APPLICATION NUMBER: US/09/438,268  
; CURRENT FILING DATE: 1999-11-10  
; EARLIER APPLICATION NUMBER: 60/107,840  
; EARLIER FILING DATE: 1998-11-10  
; EARLIER APPLICATION NUMBER: 60/123,651  
; EARLIER FILING DATE: 1999-03-10  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 8179  
; TYPE: DNA  
; ORGANISM: Virus  
US-09-438-268-5

Alignment Scores:  
Pred. No.: 4.91e-275 Length: 8179  
Score: 2518.00 Matches: 458  
Percent Similarity: 90.65% Conservative: 27  
Best Local Similarity: 85.61% Mismatches: 48  
Query Match: 86.65% Indels: 2  
DB: 4 Gaps: 2

US-09-807-802A-17 (1-534) x US-09-438-268-5 (1-8179)

Qy 1 MetAlaSerGlyGlyGlyAlaPrometAlaAspAsnGluGlyAlaAspGlyValGly 20

Db 2739 ATGGCTACAGGCAGTGGCGCACCAATGGCAGACAAATAACAGGGCGCCGACGGAGTGGGT 2798  
QY 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40  
Db 2799 AATTCTCCGGAATTTGGCAATTGGCATTCACATGGATGGGCGCAGAGTCATCACACC 2858  
QY 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
Db 2859 AGCACCGAACCCTGGGCCCTGCCACCTACAACAACCACTCTACAACAATAATTTCCAGC 2918  
QY 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80  
Db 2919 CAATCA---GGAGCCTCGAACGACATCACTACTTTGGCTACAGCACCCCTTTGGGGGTAT 2975  
QY 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100  
Db 2976 TTTGACTTCAACAGATTCCCACTGCCACTTTTCAACCACGTGACTGGCAAGACTCATCAAC 3035  
QY 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120  
Db 3036 AACAACCTGGGATTCGACCCCAAGAGACTCAACTTCAAGCTCTTTAACAATTCAGTCAAA 3095  
QY 121 GluValThrThrAsnAspGlyValThrIleAlaAsnAsnLeuThrSerThrValGln 140  
Db 3096 GAGGTACCGCAGAATGACGGTACGACGAGGATGGCAATAACCTTACCAGCACGGTTTACG 3155  
QY 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160  
Db 3156 GTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTGCTCGGGTGGCGGCACCAAGGCTGT 3215  
QY 161 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
Db 3216 CTCCCGCGCTTTCCAGCGGACGTCTTCATGTGTCCTCAGTATGGATACCTCACCCCTGAAC 3275  
QY 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
Db 3276 AACGGAAGTCAAGCGGTGGGACGCTCATCTTTTACTGCTGGAGTACTTCCCTTCGCAG 3335  
QY 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPropheHis 220  
Db 3336 ATGCTAAGGACTGGAAATAACTTCCAATTCACTATACCTTCGAGGATGTACCTTTTTCAC 3395  
QY 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
Db 3396 AGCAGTACGCTCACAGCCAGAGTTTGGATCGTTGATGAATCCTCTTATGTATGATCAGTAT 3455  
QY 241 LeuTyrTyrLeuAsnArgThrGln---AsnGlnSerGlySerAlaGlnAsnLysAspLeu 259  
Db 3456 CTGTACTACCTGAACAGAACGCAAGGAACAACCTCTGGAACAACAACCAATCACGGCTG 3515  
QY 260 LeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGly 279  
Db 3516 CTTTTTAGCCAGGCTGGGCTCAGTCTATGTCTTTGCGAGGCCAGAAATTTGGCTACCTGGG 3575  
QY 280 ProCysTyrArgGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPhe 299  
Db 3576 CCCTGTACCGGCAACAGAGACTTTTCAAAGACTGCTTAACGACAACAACAACAGTAACCTT 3635  
QY 300 ThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGly 319  
Db 3636 CCTTGGACAGCGGCCAGCAAAATATCATCTCAATGGCGCGACTCGCTGGTGAATCCAGGA 3695  
QY 320 ThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMet 339  
Db 3696 CCAGCTATGGCCAGTCAAGGACGATGAAGAAAAATTTTCCCTATGCACGGCAATCTA 3755  
QY 340 IlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThr 359  
Db 3756 ATATTGGCAAGAGGACAACGCAAGTAACGCAGAAATTAGATAATGTAAATGATTACG 3815  
QY 360 AspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAla 379

Db 3816 GATGAAGAAGAGATTCTGTACCACCAATCCTCTGTGGCAACAGAGCAGTATGGAACACTGTGGCA 3875  
QY 380 ValAsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAla 399  
Db 3876 AATAACTTGCAGAGCTCAAATACAGCTCCCAAGACTGGAACACTGTCAATCATCAGGGGGCC 3935  
QY 400 LeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLys 419  
Db 3936 TTACCTGGCATGGTGTGGCAAGATCGTGACGTGTACCTTCAAGGACCTATCTGGGCAAAAG 3995  
QY 420 IleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLys 439  
Db 3996 ATTCTCACACGGATGGACACTTTTCATCCTTCTCTCTGTATGGAGGCTTTGGACTGAAA 4055  
QY 440 AsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGlu 459  
Db 4056 CATCGCCTCTCAAATCATGATCAAAAATACTCCGGTACTCTCGGAATCCTTCGACCACC 4115  
QY 460 PheSerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerVal 479  
Db 4116 TTCAGTGGCGCAAGTTTGTCTTCTTCATCACACAGTACTCCACGGGACAGGTCAGCGTG 4175  
QY 480 GluIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyr 499  
Db 4176 GAGATCGAGTGGGAGCTGCAGAAAGGAAACAGCAAAACGCTGGAATCCCGAAATTCAGTAC 4235  
QY 500 ThrSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyr 519  
Db 4236 ACTTCCAACACTACAACAAGTCTGTTAATCGTGGACTTACCGTGGATACTAATGGCGTGTAT 4295  
QY 520 ThrGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
Db 4296 TCAGAGCCTCGCCCCCATTTGGCACCAAGATACCTGACTCGTAATCTG 4340

RESULT 2

US-09-770-315-4  
; Sequence 4, Application US/09770315  
; Patent No. 6429001  
; GENERAL INFORMATION:  
; APPLICANT: Chiron Corporation  
; TITLE OF INVENTION: Recombinant AAV Packaging Systems  
; FILE REFERENCE: 20263-501  
; CURRENT APPLICATION NUMBER: US/09/770,315  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: US 60/178,536  
; PRIOR FILING DATE: 2000-01-26  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 4072  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: recombinant DNA  
US-09-770-315-4

Alignment Scores:  
Pred. No.: 2.16e-271 Length: 4072  
Score: 2481.50 Matches: 445  
Percent Similarity: 90.26% Conservative: 37  
Best Local Similarity: 83.33% Mismatches: 51  
Query Match: 85.39% Indels: 1  
DB: 4 Gaps: 1

US-09-807-802A-17 (1-534) x US-09-770-315-4 (1-4072)

QY 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20  
Db 2090 ATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGCGCGCAGGAGTGGGT 2149  
QY 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40  
Db 2150 AATTCTCTCGGAAATTTGGCATTTCCGATTCACATGGATGGGCGCAGAGTCAATCACACCACC 2209



QY 41 SerThrArgThrTyrAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
Db 2210 AGCACCGAACCTGGGCCCTGCCACCTACAACAACCACCTCTACAACAATAATTCAGC 2269

QY 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTyrGlyTyr 80  
Db 2270 CAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTGGGGGTAT 2326

QY 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTyrGlnArgLeuIleAsn 100  
Db 2327 TTTGACTTCAACAGATTCCACTGCCACTTTTCACCAACGTGACTGGCAAGACTCATCAAC 2386

QY 101 AsnAsnTyrGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120  
Db 2387 AACAACTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAAACATTTCAAGTCAAA 2446

QY 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140  
Db 2447 GAGGTCACGCAGAATGACCGGTACGACGACGATTGCCAATAAACCTTACCAGCACGGTTCAG 2506

QY 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160  
Db 2507 GTGTTTACTGACTCGGAGTACCGACTCCCGTACGTCCTCGCTCGCTCGGCGCATCAAGGATGC 2566

QY 161 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
Db 2567 CTCCCGCCGTTCCAGCAGACGCTCTTCATGGTGGCCACAGTATGGATACCTCACCCCTGAAC 2626

QY 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
Db 2627 AACGGGAGTCAGGACGATAGGACGCTCTTCATTTTACTGCTGGAGTACTTTTCCTTCAG 2686

QY 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis 220  
Db 2687 ATGTCGCTACCGGAAACAACCTTACCTTCAGCTACACTTTTGGAGACGTTCTCTTTCCAC 2746

QY 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
Db 2747 AGCAGTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGACCATAC 2806

QY 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260  
Db 2807 CTGTATTACTTGGAGCAGAACAAACACTCCAAGTGGAAACCAACACGACGTCAGTCAAGGCTTCAG 2866

QY 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTyrLeuProGlyPro 280  
Db 2867 TTTTCTCAGGCCGGAGGTGACATTTCGGGACCACTTAGGAATGGCTTCTCTCGGACCC 2926

QY 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
Db 2927 TGTACCGCCAGCAGCGAGTATCAAAAGACATCTGGGATAAACAACAGTGAATACTCG 2986

QY 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320  
Db 2987 TGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGCCCG 3046

QY 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340  
Db 3047 GCCATGGCAAGCCACAAGGACGATGAAGAAAGTTTTTTTCTCAGAGCGGGGTTCTCATC 3106

QY 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360  
Db 3107 TTTGGGAAGCAAGGCTCAGAGAAAAACAATGTGGACATTTGAAAAGGTCTATGATTACAGAC 3166

QY 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380  
Db 3167 GAAGAGGAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTATGGTTCTGTATCTACC 3226

QY 381 AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400  
Db 3227 AACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAACACACAAAGGCGTTCTT 3286

QY 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420  
Db 3287 CCAGGCATGGTCTGGCAGGACAGAGATGTGTACTTCAGGGGCCCATCTGGGCAAGATT 3346

QY 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440  
Db 3347 CCACACACGGACGGACATTTTCACCCCTCTCCCTCATGGGTGGATTCCGACTTAACACAC 3406

QY 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460  
Db 3407 CCTCCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAATCCTTCGACCACCTTC 3466

QY 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480  
Db 3467 AGTGGGCAAGTTTGTCTCTTCATCATCACAGTACTCCACGGGACAGGTCAGCGTGGAG 3526

QY 481 IleGluTyrGluLeuGlnLysGluAsnSerLysArgTyrAsnProGluValGlnTyrThr 500  
Db 3527 ATCGAGTGGGAGCTGCAGAAGGAAACAGCAACGCTGGAATCCCGAAATTCAGTACACT 3586

QY 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520  
Db 3587 TCCAACACTACAACAAGTCTGTTAATGTGGACTTACTGTGGACACTAATGGCGGTATTCA 3646

QY 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
Db 3647 GAGCCTCGCCCAATGGCACCAGATACCTGACTCGTAATCTG 3688

RESULT 3  
US-09-770-315-3  
; Sequence 3, Application US/09770315  
; Patent No. 6429001  
; GENERAL INFORMATION:  
; APPLICANT: Chiron Corporation  
; TITLE OF INVENTION: Recombinant AAV Packaging Systems  
; FILE REFERENCE: 20263-501  
; CURRENT APPLICATION NUMBER: US/09/770,315  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: US 60/178,536  
; PRIOR FILING DATE: 2000-01-26  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 7557  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: recombinant DNA  
US-09-770-315-3

Alignment Scores:  
Pred. No.: 6,09e-271 Length: 7557  
Score: 2481.50 Matches: 445  
Percent Similarity: 90.26% Conservative: 37  
Best Local Similarity: 83.33% Mismatches: 51  
Query Match: 85.39% Indels: 1  
DB: 4 Gaps: 1

US-09-807-802A-17 (1-534) x US-09-770-315-3 (1-7557)

QY 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20  
Db 2837 ATGGCTACAGGAGTGGCGCACCAATGGCAGACAATAACGAGGGCGCGAGTGGGT 2896

QY 21 AsnAlaSerGlyAsnTyrHisCysAspSerThrTyrLeuGlyAspArgValIleThrThr 40  
Db 2897 AATTCTCGGGAATTTGGCATTTCCGATTCCCATGGATGGGCGACAGAGTCATCACACC 2956

QY 41 SerThrArgThrTyrAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
Db 2957 AGCACCAGAACCTGGGCCCTGCCACCTACACAACACCCTCTACAAACAATAATTTCCAGC 3016

QY 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTyrGlyTyr 80

|    |      |   |      |
|----|------|---|------|
| Db | 3017 | CAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTGGGGGTAT   | 3073 |
| QY | 81   | PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn  | 100  |
| Db | 3074 | TTTGACTTCAACAGATCCACTGCCACTTTTACCACGTGACTGGCAAGACTCATCAAC     | 3133 |
| QY | 101  | AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys  | 120  |
| Db | 3134 | AACAACTGGGGATTCCGACCCAAAGAGACTCAACTTCAAGCTCTTTAACATTCAAGTCAAA | 3193 |
| QY | 121  | GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln  | 140  |
| Db | 3194 | GAGGTCAAGCAGAAATGACGGTACGACGACGATTGCCAATAACCTTACCAGCACGGTTCAG | 3253 |
| QY | 141  | ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys  | 160  |
| Db | 3254 | GTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTCCTCGGCTCGGGCGCATCAAGGATGC | 3313 |
| QY | 161  | LeuProPropheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn  | 180  |
| Db | 3314 | CTCCCGCGTTCCAGCAGACGCTCTTCATGGTGGCCACAGTATGGATACCTCACCTGAAC   | 3373 |
| QY | 181  | AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln  | 200  |
| Db | 3374 | AACGGGAGTCAAGGCAGTAGGACGCTCTTCATTTTACTGCTGGAGTACTTTCTCTCTCAG  | 3433 |
| QY | 201  | MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis  | 220  |
| Db | 3434 | ATGCTGCGTACCGGAAACAACCTTTACCTTCAGCTACACTTTTGAGGACGTTCTCTTCCAC | 3493 |
| QY | 221  | SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr  | 240  |
| Db | 3494 | AGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGACCAAGTAC | 3553 |
| QY | 241  | LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu  | 260  |
| Db | 3554 | CTGTATTACTTGAGCAGACAACAACACTCCAAGTGGAAACCACCACGCAAGGCTTCAG    | 3613 |
| QY | 261  | PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro  | 280  |
| Db | 3614 | TTTTCTCAGGCCGAGCGAGTGACATTCGGGACCAGTCTAGGAACTGGCTTCTTGACCC    | 3673 |
| QY | 281  | CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr  | 300  |
| Db | 3674 | TGTTACCGCCAGCAGCGAGTATCAAAAGACATCTGCGGATAACAACAACAGTGAATACTCG | 3733 |
| QY | 301  | TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr  | 320  |
| Db | 3734 | TGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGCCG   | 3793 |
| QY | 321  | AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle  | 340  |
| Db | 3794 | GCCATGGCAAGCCACAAGGACGATGAAGAAAGTTTTTCTCTCAGAGCGGGGTTCTCATC   | 3853 |
| QY | 341  | PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp  | 360  |
| Db | 3854 | TTTGGGAAGCAAGGCTCAGAGAAAACAATGTGGACATTGAAAAGGTCTATTGATTACAGAC | 3913 |
| QY | 361  | GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal  | 380  |
| Db | 3914 | GAAGAGGAAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTATGGTTCTGTATCTACC  | 3973 |
| QY | 381  | AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu  | 400  |
| Db | 3974 | AACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAACACACACAGGCGTCTT  | 4033 |
| QY | 401  | ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle  | 420  |
| Db | 4034 | CCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGCAAGATT   | 4093 |
| QY | 421  | ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn  | 440  |

|            |      |   |      |
|------------|------|---|------|
| D <b>b</b> | 4094 | CCACACACGACGGACATTTTTCACCCCTCTCCCCCTCATGGTGGATTCGGAGCTTAAACAC | 4153 |
| Q <b>y</b> | 441  | ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe  | 460  |
| D <b>b</b> | 4154 | CCTCCTCCACAGATTCTCATCAAGAAACACCCCGGTACCTGCGAATCCTTCGACCACCTTC | 4213 |
| Q <b>y</b> | 461  | SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu  | 480  |
| D <b>b</b> | 4214 | AGTGGCGCAAGTTTGCTTCTTCATCACACAGTACTCCACGGGACAGGTACCGTGGAG     | 4273 |
| Q <b>y</b> | 481  | IleGluTyrGluLeuGlnLysGluAsnSerLysArgTyrAsnProGluValGlnTyrThr  | 500  |
| D <b>b</b> | 4274 | ATCGAGTGGGAGCTGCAGAAGGAAACAGCAAACGCTGGAATCCCGAAATTTCAGTACACT  | 4333 |
| Q <b>y</b> | 501  | SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr  | 520  |
| D <b>b</b> | 4334 | TCCAACTACAACAAGTCTGTTAATGTGGACTTACTGTGGACACTAATGGCGTGATTCA    | 4393 |
| Q <b>y</b> | 521  | GluProArgProIleGlyThrArgTyrLeuThrArgProLeu                    | 534  |
| D <b>b</b> | 4394 | GAGCCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTG                    | 4435 |

## RESULT 4

```

US-09-770-315-2
; Sequence 2, Application US/09770315
; Patent No. 6429001
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; TITLE OF INVENTION: Recombinant AAV Packaging Systems
; FILE REFERENCE: 20263-501
; CURRENT APPLICATION NUMBER: US/09/770,315
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,536
; PRIOR FILING DATE: 2000-01-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 8698
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: recombinant DNA
US-09-770-315-2

```

|                        |           |               |  |  |      |
|------------------------|-----------|---------------|--|--|------|
| Alignment Scores:      |           |               |  |  |      |
| Pred. No.:             | 7.71e-271 | Length:       |  |  | 8698 |
| Score:                 | 2481.50   | Matches:      |  |  | 445  |
| Percent Similarity:    | 90.26%    | Conservative: |  |  | 37   |
| Best Local Similarity: | 83.33%    | Mismatches:   |  |  | 51   |
| Query Match:           | 85.39%    | Indels:       |  |  | 1    |
| DB:                    | 4         | Gaps:         |  |  | 1    |

US-09-807-802A-17 (1-534) x US-09-770-315-2 (1-8698)

|    |      |   |      |
|----|------|---|------|
| QY | 1    | MetAlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly  | 20   |
| DB | 2809 | ATGGCTACAGGCAGTCGGCGCACCAATGGCAGACATAAACGAGGGCGCCGACGGAGTGGGT | 2868 |
| QY | 21   | AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr  | 40   |
| DB | 2869 | AATTCCTCGGGAATGGCATTGGGATTCACATGGATGGCGCACAGAGTCATCACCACC     | 2928 |
| QY | 41   | SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer  | 60   |
| DB | 2929 | AGCACCCGAACTGGGGCCCTGCCCCCTACAACAACACCACCTCTACAAACAAATTTCCAGC | 2988 |
| QY | 61   | AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr  | 80   |
| DB | 2989 | CAATCA--GGAGCCTCGAACGACAAATCACTACTTTGGCTACAGCACCCCTTGGGGGGTAT | 3045 |
| QY | 81   | PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn  | 100  |
| DB | 3046 | TTTGACTTCAACAGATTCACCTGCCACTTTTACCACCGTGACTGGCAAGACTCATCAAC   | 3105 |

Qy 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPhelLysLeuPheAsnIleGlnValLys 120  
Db 3106 AACAACTGGGGATTCCGACCCAAAGAGACTCAACTTCAAGCTCTTTAAACATTCAAGTCAA 3165  
Qy 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140  
Db 3166 GAGGTCAGCAGAAATGACGGTACGACGAGATGGCCAAATAACCTTACCAGCAGCGTTTCAG 3225  
Qy 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160  
Db 3226 GTGTTTACTGACTCGGAGTACGAGCTCCCGTACGTCCTCGGCTCGGCGCATCAAGGATGC 3285  
Qy 161 LeuProPropheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
Db 3286 CTCCCGCGGTTCCAGCAGAGAGCTCTTCATGGTGCCACAGTATGGATACCTCACCGCTGAAC 3345  
Qy 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
Db 3346 AACGGGAGTCAGGCAGTAGGACGCTCTTCATTTACTGCTGGAGTACTTTCCTTCTCAG 3405  
Qy 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis 220  
Db 3406 ATGCTGCGTACCGGAAACAACCTTACCTTCAGCTACACTTTTGAGGACGTTCTCTTCCAC 3465  
Qy 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
Db 3466 AGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGACCATC 3525  
Qy 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260  
Db 3526 CTGTATTACTTGAGCAGAACAAACACTCTCAAGTGGAAACCACCGCAGTCGAGCTTCAG 3585  
Qy 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280  
Db 3586 TTTTCTCAGGCGGAGCGAGTGACATTCGGGACCAGTCTAGGAACCTGGCTTCCTGGACCC 3645  
Qy 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
Db 3646 TGTATCCGCGCAGCAGCGAGTATCAAAAGACATCTGCGGATAACAACAACAGTGAATACTCG 3705  
Qy 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320  
Db 3706 TGGACTGGAGTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGCCCCG 3765  
Qy 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340  
Db 3766 GCCATGGCAAGCCACAAAGGACGATGAAGAAAGTTTTTCTCAGAGCGGGTTCTCATC 3825  
Qy 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360  
Db 3826 TTTGGGAAGCAAGGCTCAGAGAAACAATAATGTGGACATTGAAGAAGGTCTATGATTACAGAC 3885  
Qy 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380  
Db 3886 GAAGAGAAATCAGGACAACCAATCCCGTGGCTACGGAGCAGTATGGTTCTGTATCTACC 3945  
Qy 381 AsnPheGlnSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400  
Db 3946 AACCTCCAGAGAGGCAACAGACAGACAGCTACCGCAGATGTCAACACACAAAGCGTTCTT 4005  
Qy 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420  
Db 4006 CCAGGATGGTCTGGCAGGACAGAGATGTACCTTCAGGGGCCCCATCTGGGCAAGATT 4065  
Qy 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440  
Db 4066 CCACACAGGACGGACATTTTACCCCTCTCCCTCATGGGTGGATTTCGGACTTAAACAC 4125  
Qy 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460  
Db 4126 CCTCCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAATCCTTCGACCACCTTC 4185

Qy 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGlu 480  
Db 4186 AGTCGGGCAAAAGTTTGTCTTCCATCACACAGTACTCCACGGGACAGTCCAGCGTGGAG 4245  
Qy 481 IleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrThr 500  
Db 4246 ATCGAGTGGGAGCTGCAGAGGAAACAGCAAAACGCTGGAATCCCGAAATTCAGTACACT 4305  
Qy 501 SerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThr 520  
Db 4306 TCCAACTACACAAGTCTGTTAATGTGGACTTTACTGTGACACTAATGGCGGTATTCA 4365  
Qy 521 GluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
Db 4366 GAGCCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTG 4407

RESULT 5

US-08-254-358-1  
; Sequence 1, Application US/08254358  
; Patent No. 5658785  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Philip R.  
; TITLE OF INVENTION: Adeno-Associated Virus Materials and  
; TITLE OF INVENTION: Methods  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 S. Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/254,358  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5658785and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 31975  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4680 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-254-358-1

Alignment Scores:  
Pred. No.: 1.63e-266 Length: 4680  
Score: 2439.50 Matches: 441  
Percent Similarity: 89.16% Conservative: 36  
Best Local Similarity: 82.43% Mismatches: 56  
Query Match: 83.95% Indels: 2  
DB: 1 Gaps: 1

US-09-807-802A-17 (1-534) x US-08-254-358-1 (1-4680)

Qy 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20  
Db 2809 ATGGCTACAGGCAGTGGCGCACCAATGGCAGACAATAACGAGGGCGCGAGTGGGT 2868  
Qy 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40



Db 2869 AATTCTCCGAAATTGGCATTTGGCATTTCCACATGGATGGGCGACAGAGTCAACACC 2928  
Qy 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
Db 2929 AGCACCGGAACCTGGGCCCTGCCACCTTACAAACACCACTCTACAAACAAATTTCCAGC 2988  
Qy 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTyrGlyTyr 80  
Db 2989 CAATCA---GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTTGGGGGTAT 3045  
Qy 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100  
Db 3046 TTTGACTTCAACAGATTCCACTGCCACTTTTCACCACGTGCTGGCAAGACTCATCAAC 3105  
Qy 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120  
Db 3106 AACAACTGGGATTCGACCCCAAGAGACTCAACTTCAAGCTCTTAAACATTTCAAGTCAA 3165  
Qy 121 GluValThrThrAsnAspGlyValThrIleAlaAsnAsnLeuThrSerThrValGln 140  
Db 3166 GAGGTACGCAGAAATGACGGTACGACGAGATTGCCAATAACCTTACCAGCACGGTTTACG 3225  
Qy 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160  
Db 3226 GTGTTTACTACTCGGAGTACCAGCTCCGCTACGCTCTCGGCTCGGCGCATCAAGGATGC 3285  
Qy 161 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
Db 3286 CTCCCGCCGTTCCAGCAGACAGCTTTCATGGTGCCACAGTATGGATACCTCACCCCTGAAC 3345  
Qy 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
Db 3346 AACGGGAGTCAGCAGTAGGACGCTCTTCAATTTACTGCTGGAGTACTTTCTCTTCTCAG 3405  
Qy 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis 220  
Db 3406 ATGCTGCGTACCGGAAACAACCTTTACCTTCAGCTACACATTTTGAGGACGTTCTCTTCCAC 3465  
Qy 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
Db 3466 AGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGACCAGTAC 3525  
Qy 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260  
Db 3526 CTGTATTACTTGACAGACAACAACACTCCAGTGAACCAACCACCGCAGTCAAGGCTTCAG 3585  
Qy 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280  
Db 3586 TTTTCTCAGGCCCGGAGCGAGTGACATTCGGGACCAAGTCTAGGAACCTGCTTCTGGACCC 3645  
Qy 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
Db 3646 TGTTACCGCCAGCAGCGAGTATCAAAGACATCTGCGGATACAAACACAGTGAATACTCG 3705  
Qy 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleLeuAsnProGlyThr 320  
Db 3706 TGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGGCCC 3765  
Qy 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetile 340  
Db 3766 GCCATGGCAAGCCCAAGGACGATGAAGAAAAGTTTTCCTCAGAGCGGGTCTCTCATC 3825  
Qy 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetileThrAsp 360  
Db 3826 TTTGGGAAGCAAGGCTCAGAGAAACAATGTGAACATTTGAAAGGTGATGATACAGAC 3885  
Qy 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380  
Db 3886 GAAGAGGAAATCGGAACAACCAATCCCGTGGCTACGGAGCAGTATGGTGTGTATCTACC 3945  
Qy 381 AsnPheGlnSerSerThrThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400

Db 3946 AACCTCCAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAACACACAAGGCGTTCTT 4005  
Qy 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420  
Db 4006 CCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTCTGGGCAAGATT 4065  
Qy 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440  
Db 4066 CCACACACGACGGACATTTTCAACCCCTCTCCCTCATGGGTGGATTCTGGACTTAAACAC 4125  
Qy 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460  
Db 4126 CCTCTCCACAGATTCTCATCAAGAACACCCCGGTACTCTCGAATCCTTCGACCACTTC 4185  
Qy 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGln-ValSerValG1 480  
Db 4186 AGTGGGCAAGTTTGTCTTCTTCTCATCACACAGTACTCCACGGGACACGGTCAGCGTGA 4245  
Qy 480 uIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrTh 500  
Db 4246 GATCGAGTGGGAGCTCGAAGGAAACACAGCAACGCTGGAATCCCGAAATTCAGTACAC 4305  
Qy 500 rSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrTh 520  
Db 4306 TTCCAACACTACAACAAGTCTGTTAATCGTGGACTTACCGTGGATACTAATGGCGTGTATTC 4365  
Qy 520 rGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
Db 4366 AGAGCCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTG 4408

RESULT 6  
US-08-475-391-1  
; Sequence 1, Application US/08475391  
; Patent No. 5786211  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Philip R.  
; TITLE OF INVENTION: Adeno-Associated Virus Materials and  
; TITLE OF INVENTION: Methods  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 S. Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/475,391  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/254,358  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5786211and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 31975  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4680 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)

US-08-475-391-1

Alignment Scores:

Pred. No.: 1.63e-266 Length: 4680  
Score: 2439.50 Matches: 441  
Percent Similarity: 89.16% Conservative: 36  
Best Local Similarity: 82.43% Mismatches: 56  
Query Match: 83.95% Indels: 2  
DB: 1 Gaps: 1

US-09-807-802A-17 (1-534) x US-08-475-391-1 (1-4680)

Qy 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnGluGlyAlaAspGlyValGly 20  
Db 2809 ATGGCTACAGGAGTGGCGCACCAATGGCAGACAATAACGAGGGCGCCGAGTGGT 2868  
Qy 21 AsnAlaSerGlyAsnTyrHisCysAspSerThrTyrLeuGlyAspArgValIleThr 40  
Db 2869 AATTCTCTCCGAAATTGGCATTTGGCATTTCCACATGGATGGCGCAGAGTCATCACCACC 2928  
Qy 41 SerThrArgThrTyrAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
Db 2929 AGCACCCGAACCTGGGCCCTGCCACCTACAACAACCACTCTACAACAATAATTTCCAGC 2988  
Qy 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTyrGlyTyr 80  
Db 2989 CAATCA--GGAGCCTCGAACGACAAATCACTACTTTGGCTACAGCACCCCTTGGGGTAT 3045  
Qy 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTyrGlnArgLeuIleAsn 100  
Db 3046 TTTGACTTCAACAGATTCCCACTGCCACTTTTCCACCAGTGAAGTGGCAAGACTCATCAAC 3105  
Qy 101 AsnAsnTyrGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120  
Db 3106 AACAACTGGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAAACATTCAAGTCAAA 3165  
Qy 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140  
Db 3166 GAGGTACGCAGAAATGACGGTACGACGACGATTGCCAATAACCTTACCAGCACGGTTTACG 3225  
Qy 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160  
Db 3226 GTGTTTACTGACTCGGAGTACCAAGTCCCGTACGTCTCGGTGGCGCATCAAGGATGC 3285  
Qy 161 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
Db 3286 CTCCCGCGGTTCCAGCAGACGTCTTCATGGTGGCCACAGTATGGATACCTCACCCCTGAAC 3345  
Qy 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
Db 3346 AACGGAGTCAGGCAGTAGGACGCTCTTCACTTTTACTGCTGGAGTACTTTCTCTCTCAG 3405  
Qy 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis 220  
Db 3406 ATGCTGCTACCGGAAACAACCTTTACCTTCAGCTACACTTTTGGAGCGTTCTCTTCCAC 3465  
Qy 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
Db 3466 AGCAGTACGCTCACAGCCAGAGTCTGGACCGTCTCATGTATCTCTCATCGACCACTAC 3525  
Qy 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260  
Db 3526 CTGTATTACTTGAGCAGAAACAACACTCCAAGTGGAAACCAACGACGAGTCAAGGCTTCAG 3585  
Qy 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTyrLeuProGlyPro 280  
Db 3586 TTTTCTCAGGCCGAGGAGTGACATTCGGGACCACTAGTGTAGGAACCTGGCTTCTCTGGACCC 3645  
Qy 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
Db 3646 TGTTACCGCCAGCAGGATATCAAAAGACATCTCGGGAATAACAACAACAGTGAATACTCG 3705  
Qy 301 TrrThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320

Db 3706 TGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGCCCC 3765  
Qy 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340  
Db 3766 GCCATGGCAAGCCCAAGGACGATGAAGAAAAGTTTTCCTCAGAGCGGGTTCTCATC 3825  
Qy 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360  
Db 3826 TTTGGGAAGCAAGGCTCAGAGAAAACAATGTGAACATTGAAAAGGTTCATGATTACAGAC 3885  
Qy 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380  
Db 3886 GAAGAGGAAATCGGAACAACCAATCCCCGTGGCTACGGAGCAGTATGGTTCTGTATCTACC 3945  
Qy 381 AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400  
Db 3946 AACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAACACACAAGCGCTTCTT 4005  
Qy 401 ProGlyMetValTrrGlnAspArgAspValTyrLeuGlnGlyProIleTrrAlaLysIle 420  
Db 4006 CCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGCAAGATT 4065  
Qy 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440  
Db 4066 CCACACACGGACGGACATTTTCAACCCCTCTCCCTCATGGGTGGATTCCGACTTAAACAC 4125  
Qy 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460  
Db 4126 CCTCTCCACAGATTCTCATCAAGAACAACCCCGGTACCTGCGAATCCTTCGACCACTTC 4185  
Qy 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGln-ValSerValGln 480  
Db 4186 AGTGGGGCAAAGTTTGTCTTCTTCATCATCACAGTACTCCACGGGACACGGTCAGCGTGA 4245  
Qy 480 uileGluTrrGluLeuGlnLysGluAsnSerLysArgTrrAsnProGluValGlnTyrTh 500  
Db 4246 GATCGAGTGGGAGCTGCAGAAAGGAAAACAGCAAACGCTGGAATCCGAAATTCAGTACAC 4305  
Qy 500 rSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrTh 520  
Db 4306 TTCCAACCTACAAACAAGTCTGTTAATCGTGGACTTACCGTGGATACTAATGGCGTGTATTC 4365  
Qy 520 rGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
Db 4366 AGAGCCTCGCCCCCATTTGGCACCAGATACCTGACTCGTAATCTG 4408

RESULT 7

US-08-709-609-1  
; Sequence 1, Application US/08709609  
; Patent No. 5858775  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Philip R.  
; TITLE OF INVENTION: Adeno-Associated Virus Materials and  
; TITLE OF INVENTION: Methods  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 S. Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/709,609  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:

NAME: No. 5858775and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 31975  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4680 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-709-609-1.

Alignment Scores:  
Pred. No.: 1.63e-266 Length: 4680  
Score: 2439.50 Matches: 441  
Percent Similarity: 89.16% Conservative: 36  
Best Local Similarity: 82.43% Mismatches: 56  
Query Match: 83.95% Indels: 2  
DB: 2 Gaps: 1

US-09-807-802A-17 (1-534) x US-08-709-609-1 (1-4680)

QY 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20  
DB 2809 ATGGCTACAGGAGTGGCGCACCAATGGCAGACAATAACGAGGGCGCGAGTGGGT 2868  
QY 21 AsnAlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThr 40  
DB 2869 AATTCTCCGGAAATTGGCATTTCCGATTTCCACATGGATGGCGACAGAGTCATACCACC 2928  
QY 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
DB 2929 AGCACCGGAACCTGGGCGCTGCCACCTACAAACACCACTCTACAAACAATAATTCCAGC 2988  
QY 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80  
DB 2989 CAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTGGGGGTAT 3045  
QY 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100  
DB 3046 TTTGACTTCAACAGATTCCACTGCCACTTTTTCACCACGTGACTGGCAAGACTCATCAAC 3105  
QY 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120  
DB 3106 AACAACTGGGGATTCCGCCCAAGAGACTCAACTTCAAGCTCTTTAACATTCAAGTCAAA 3165  
QY 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140  
DB 3166 GAGTCCAGCAGAAATACGGTACGACGACGATTTGCCAATAACCTTACCAGCAGGTTTCAG 3225  
QY 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160  
DB 3226 GTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTCTCGGCTCGGCGCATCAAGGATGC 3285  
QY 161 LeuProProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180  
DB 3286 CTCCCGCGGTTCCCGACGACAGCTCTTCATGGTGCCACAGTATGGATACCTCACCTGAAC 3345  
QY 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
DB 3346 AACGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCTGGAGTACTTTCTCTCTCAG 3405  
QY 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHis 220  
DB 3406 ATGTCGCTACCGGAACAACCTTACCTTCAGCTACACTTTTGAGGACGTTCTCTTCCAC 3465  
QY 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
DB 3466 AGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGACCATAC 3525

QY 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260  
DB 3526 CTGTATTACTTGAGCAGAAACAACACTCCAAGTGGAAACCAACCGCAGTCAAGGCTTCAG 3585  
QY 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280  
DB 3586 TTTTCTCAGGCGGAGCGAGTACATTCGGGACCAGTCTAGGAACTGGCTTCTCTGGACCC 3645  
QY 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
DB 3646 TGTACCGCCAGCAGCGAGTATCAAAGACATCTCGCGATAACAACAACAGTGAATACTCG 3705  
QY 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320  
DB 3706 TGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGGCCC 3765  
QY 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340  
DB 3766 GCCATGGCAAGCCACAAGGACGATGAAGAAAGTTTTTCTCAGAGCGGGTTCTCATC 3825  
QY 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360  
DB 3826 TTTGGGAAGCAAGGCTCAGAGAAAACAAATGTGAACATTGAAAAGGTCTATGATTACAGAC 3885  
QY 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380  
DB 3886 GAAGAGGAAATCGGAACAACCAATCCCGTGGCTACGGAGCAGTATGTTCTGTATCTACC 3945  
QY 381 AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400  
DB 3946 AACCTCCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAACACACACAAGCGTCTT 4005  
QY 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420  
DB 4006 CCAGGATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTGGCAAGATT 4065  
QY 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440  
DB 4066 CCACACAGGACGGACATTTTCAACCCCTCTCCCTCATGTGGTGGATTCTGGACTTAAACAC 4125  
QY 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460  
DB 4126 CCTCCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAATCCTTCGACCACCTTC 4185  
QY 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGln-ValSerValG 480  
DB 4186 AGTGGCGCAAGATTGCTTCTTCTCATCACACAGTACTCCACGGGACACGGTTCAGCGTGA 4245  
QY 480 uIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrTh 500  
DB 4246 GATCGAGTGGAGCTGCAGAGGAAAACAGCAACCGTGGAAATCCGAAATTCAGTACAC 4305  
QY 500 rSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrTh 520  
DB 4306 TTCCAACCTACAACAAGTCTGTTAATCGTGGACTTACCGTGGATATAATGGCGTGTATTC 4365  
QY 520 rGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
DB 4366 AGAGCCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTG 4408

RESULT 8  
PCT-US95-07178-1  
; Sequence 1, Application PC/TUS9507178  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Philip R.  
; TITLE OF INVENTION: Adeno-Associated Virus Materials and  
; TITLE OF INVENTION: Methods  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 S. Wacker Drive  
; CITY: Chicago



STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/07178  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Noland, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 31975  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4680 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
PCT-US95-07178-1

Alignment Scores:  
Pred. No.: 1.63e-266 Length: 4680  
Score: 2439.50 Matches: 441  
Percent Similarity: 89.16% Conservative: 36  
Best Local Similarity: 82.43% Mismatches: 56  
Query Match: 83.95% Indels: 2  
DB: 5 Gaps: 1

US-09-807-802A-17 (1-534) x PCT-US95-07178-1 (1-4680)

QY 1 MetAlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGly 20  
Db 2809 ATGGCTACAGGAGTGGCGCACCAATGGCAGACAATAACGAGGGCGCGAGTGGGT 2868  
QY 21 AsnAlaSerGlyAsnTrpHisCysAspSerTrpLeuGlyAspArgValIleThrThr 40  
Db 2869 AATTCTCCGGAATTGGCATTTGCCATTCCACATGGATGGCGACAGAGTTCATCACCACC 2928  
QY 41 SerThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSer 60  
Db 2929 AGCACCCGAACCTGGGCCCTGCCACCTACACCAACCACTCTACAAACAATTTCCAGC 2988  
QY 61 AlaSerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyr 80  
Db 2989 CAATCA--GGAGCCTCGAACGACAATCACTACTTTGGCTACAGCACCCCTTGGGGGTAT 3045  
QY 81 PheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsn 100  
Db 3046 TTTGACTTCAACAGATTCCACTGCCACTTTTCCACCACTGACTGGCAAGACTCATCAAC 3105  
QY 101 AsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLys 120  
Db 3106 AACAACTGGGATTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAACATTCAAGTCNAA 3165  
QY 121 GluValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGln 140  
Db 3166 GAGGTCACGCAGAAATGACCGGTACGACGAGATTGCCAATAACCTTACCAGCAGGTTTCAG 3225  
QY 141 ValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCys 160  
Db 3226 GTGTTTACTGACTCGGAGTACCAGCTCCCGTACGTCTCGGCTCGGCGCATCAAGGATGC 3285  
QY 161 LeuProPropheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsn 180

Db 3286 CTCCCGCGTTCCAGCAGACGCTCTTCATGGTGCCACAGTATGGATACCTCACCCCTGAAC 3345  
QY 181 AsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGln 200  
Db 3346 AACGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCTGGAGTACTTTCTCTCTCAG 3405  
QY 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPropheHis 220  
Db 3406 ATGCTGCGTACCGGAAACAACCTTACCTTCAGCTACACTTTTGAGGACGTTCTCTTCCAC 3465  
QY 221 SerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyr 240  
Db 3466 AGCAGCTACGCTCACAGCCAGATCTGGACCGTCTCATGAATCCTCTCATCGACCAGTAC 3525  
QY 241 LeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeu 260  
Db 3526 CTGTATTACTTGAGCAGACAAACACTCCAAGTGGAGCAACCAACAGTCAAGGCTTCAG 3585  
QY 261 PheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyPro 280  
Db 3586 TTTTCTCAGCGCGGAGTGACATTCGGGACCAGTCTAGGAACCTGGCTTCTCTGGACCC 3645  
QY 281 CysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThr 300  
Db 3646 TGTTACCGCCAGCAGCGAGTATCAAGACATCTGCGGATAACAACAACAGTGAATACTCG 3705  
QY 301 TrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThr 320  
Db 3706 TGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGGCCC 3765  
QY 321 AlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIle 340  
Db 3766 GCCATGGCAAGCCACAGGAGCATGAAGAAAAGTTTTTCTCAGAGCGGGGTCTCTCATC 3825  
QY 341 PheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAsp 360  
Db 3826 TTTGGGAAGCAAGGCTCAGAGAAAACAATGTGAACATTGAAAAGGTCTATGATTACAGAC 3885  
QY 361 GluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaVal 380  
Db 3886 GAAGAGGAATCGGAACAACCAATCCCGTGGCTACGGAGCAGTATGGTTCTGTATCTACC 3945  
QY 381 AsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeu 400  
Db 3946 AACCTCCAGAGAGGCAACAGACAGCAGTACCGCAGATGTCAACACACACAGGGCTCTT 4005  
QY 401 ProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIle 420  
Db 4006 CCAGGCATGGTCTGGCAGGACAGAGATGTGTACTCTTCAGGGGCCCATCTGGGCAAGATT 4065  
QY 421 ProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn 440  
Db 4066 CCACACAGGACGGACATTTTCAACCCCTCTCCCTCATGGTGGATTCGGACTTTAAACAC 4125  
QY 441 ProProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPhe 460  
Db 4126 CCTCCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAATCCTTCGACCACCTTC 4185  
QY 461 SerAlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGln-ValSerValGln 480  
Db 4186 AGTGGCGCAAGTTTGTCTTCTTCATCACACAGTACTCCACGGGACACGGTCAGCGTGA 4245  
QY 480 uIleGluTrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGluValGlnTyrTh 500  
Db 4246 GATCGAGTGGGAGCTGCAGAGAGGAAACAGCAACCGTGGATCCCGAAATTCAGTACAC 4305  
QY 500 rSerAsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrTh 520  
Db 4306 TTCCAACTACAACAAGTCTGTTAATCGTGGACTTACCGTGGATTAATGGCGTGTATTC 4365  
QY 520 rGluProArgProIleGlyThrArgTyrLeuThrArgProLeu 534  
Db 4366 AGAGCCTCGCCCCCATTTGGCACCAGACATACCTGACTCGTAATCTG 4408

```

RESULT 9
US-09-438-268-2
; Sequence 2, Application US/09438268
; Patent No. 6491907
; GENERAL INFORMATION:
; APPLICANT: Rabinowitz, Joseph E.
; APPLICANT: Samulski, Richard J
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 5470-186
; CURRENT APPLICATION NUMBER: US/09/438,268
; EARLIER FILING DATE: 1999-11-10
; EARLIER FILING DATE: 1998-11-10
; EARLIER FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 2
; LENGTH: 8151
; TYPE: DNA
; ORGANISM: Virus
US-09-438-268-2

Alignment Scores:
Pred. No.:      1.59e-245      Length:      8151
Score:          2258.50      Matches:      405
Percent Similarity: 85.18%      Conservative: 49
Best Local Similarity: 75.98%      Mismatches:  74
Query Match:      77.72%      Indels:       5
DB:               4          Gaps:          2

US-09-807-802A-17 (1-534) x US-09-438-268-2 (1-8151)

QY      2 AlaSerGlyGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGlyAsn 21
Db      2730 GCAGCAGCTGGCGAGCTGCAGTC--GAGGGCGGACAAAGTGCAGTGGAGTGGGTAAT 2786

QY      22 AlaSerGlyAsnTropHisCysAspSerThrTrrLeuGlyAspArgValIleThrThrSer 41
Db      2787 GCCTCGGGTGATTGGCATTCGATTCCACCTGTGTGAGGGCCACGTCAGCACCAGC 2846

QY      42 ThrArgThrTrrAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSerAla 61
Db      2847 ACCAGAACCTGGGTCTTGCCCACTTACAAACACCACCTCTACAAGCGACTCGGAGAG-- 2903

QY      62 SerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrThrProTrrGlyTyrPhe 81
Db      2904 -----AGCCTGGAGTCCAAACACCTACACGGATTCTCCACCCCTGGGATACCTTT 2954

QY      82 AspPheAsnArgPheHisCysHisPheSerProArgAspTrrPglNArgLeuIleAsnAsn 101
Db      2955 GACTTCAACCGCTTCCACTGCCACTTCTCACCACGTGACTGGCAGCGACTCATCAACAAC 3014

QY      102 AsnTrrPglPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLysGlu 121
Db      3015 AACTGGGGCATCGCAGCCCAAGCCATCGGGGTCAAAATCTTCAACATCCAGTCAAGGAG 3074

QY      122 ValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGlnVal 141
Db      3075 GTCACGACGTGCAACGGCGAGACAAACGGTGGCTTAATAACCTTACCAGCACGGTTCAGATC 3134

QY      142 PheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCysLeu 161
Db      3135 TTTCGGGACTCGTCGTACGAACTGCCGTACGTCTCGCTCGGCGCATCAAGGATGCCTC 3194

QY      162 ProPropheProAlaAspValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsnAsn 181
Db      3195 CCGCCGTTCCAGCAGACGTCTTCATGGTGGTCCACAGTATGGATACCTCACCTGAACAAC 3254

QY      182 GlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPheProSerGlnMet 201
```

```

Db      3255 GGGAGTCAGGCAGTAGGACGCTCTTCATTTTACTGCCTGGAGTACTTTCTCTTCAGATG 3314

QY      202 LeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValProPheHisSer 221
Db      3315 CTGCGTACCGGAAACAACCTTTACCTTCAGCTACACCTTTTGGAGGACGTTCTCTTCCACAGC 3374

QY      222 SerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyrLeu 241
Db      3375 AGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCCTCTCATCGACCAGTACCTG 3434

QY      242 TyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeuPhe 261
Db      3435 TATTACTTGAGCAGACAACAACACTCCAAGTGGAAACCACCACGCGAGTCAAGGCTTCAGTTT 3494

QY      262 SerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrrLeuProGlyProCys 281
Db      3495 TCTCAGGCCGAGCGAGTGACATTCGGGACCAGTCTAGGAACCTGGCTTCTCTGGACCTGT 3554

QY      282 TyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThrTrr 301
Db      3555 TACCGCCAGCAGCGAGTATCAAGACATCTCGGGATAACAACAACAGTGAATATCTCGTGG 3614

QY      302 ThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThrAla 321
Db      3615 ACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGCCCGGCC 3674

QY      322 MetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGlyValMetIlePhe 341
Db      3675 ATGCAAGCCACAAAGCAGTGAAGAAAGTTTTCCTCAGAGCGGGGTTCTCATCTTT 3734

QY      342 GlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAspGlu 361
Db      3735 GGGAAACAAGGCTCAGAGAAAACAATGTGAACATTTGAAAGGTGATGATTACAGACGAA 3794

QY      362 GluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThrValAlaValAsn 381
Db      3795 GAGGAATCGGAACAACCAATCCCGTGGTACCGGAGCAGTATGGTTCTGTATCTACCAAC 3854

QY      382 PheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeuPro 401
Db      3855 CTCAGAGAGGCAACAGACAAGCAGCTACCGCAGATGTCAACACACACAAAGCGTCTTCCA 3914

QY      402 GlyMetValTrrPglNAspArgAspValTyrLeuGlnGlyProIleTrrAlaLysIlePro 421
Db      3915 GGCATGCTCTGGCAGGACAGAGATGTACCTTCAGGGGCCCATCTGGGCAAGATTCCA 3974

QY      422 HisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsnPro 441
Db      3975 CACACGGACGGACATTTTCACCCCTCTCCCTCATGGGTGGATTGGGACTTAAACACCCCT 4034

QY      442 ProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPheSer 461
Db      4035 CCTCCACAGATTCTCATCAAGAACACACCCCGGTACCTCGGAATCCTTCGACCACCTTCAGT 4094

QY      462 AlaThrLysPheAlaSerPheIleThrGlnTyrSerThrGlyGlnValSerValGluIle 481
Db      4095 GCGGCAAAAGTTTGCTTCTTCATCACACAGTACTCCACGGGACAGGTGAGGATC 4154

QY      482 GluTrrPglLeuGlnLysGluAsnSerLysArgTrrAsnProGluValGlnTyrThrSer 501
Db      4155 GAGTGGGAGCTGCAGAGGAAACAGCAACCGTGGAAATCCCGAAATTCAGTACACTTCC 4214

QY      502 AsnTyrAlaLysSerAlaAsnValAspPheThrValAspAsnAsnGlyLeuTyrThrGlu 521
Db      4215 AACTACAACAAGTCTGTTAATCGTGGACTTACCGTGGATACTAATGGCGTGTATTTCAGAG 4274

QY      522 ProArgProIleGlyThrArgTyrLeuThrArgProLeu 534
Db      4275 CCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTG 4313
```

RESULT 10

US-09-532-594B-19

```
; Sequence 19, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 1617
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: synthetic construct
; NAME/KEY: misc_feature
; LOCATION: 162
; OTHER INFORMATION: n = g, a, c or t(u)
; NAME/KEY: misc_feature
; OTHER INFORMATION: AAV4 capsid protein vp3 gene
US-09-532-594B-19

Alignment Scores:
Pred. No.:      4.19e-182      Length:      1617
Score:          1690.50      Matches:      325
Percent Similarity: 71.30%      Conservative: 65
Best Local Similarity: 59.41%      Mismatches: 132
Query Match:    58.17%      Indels:      25
DB:              4          Gaps:        9

US-09-807-802A-17 (1-534) x US-09-532-594B-19 (1-1617)

Qy      2 AlaSerGlyGlyGlyAlaProMetAlaAspAsnGluGlyAlaAspGlyValGlyAsn 21
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      7 GCAGCAGCTGGCGGAGCTGCAGTC--GAGGGGGGACAAAGGTGCGGATGGAGTGGTAAT 63

Qy      22 AlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThrSer 41
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      64 GCCTCGGSGTATTGGCATTGGCATTCCACCTGGTCTGAGGGGCCACGTCACGACCACCGC 123

Qy      42 ThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSerAla 61
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      124 ACCAGAACCTGGGTCTTGCCCACTTACCAACACCACCTNTACAAGCGACTCGGAGAG--- 180

Qy      62 SerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyrPhe 81
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      181 -----AGCTGCAGTCCCAACACTACAACGGATTCTCCACCCCTGGGATACTTT 231

Qy      82 AspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsnAsn 101
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      232 GACTTCAACCGCTTCCACTGCCACTTCTCACCACGTCGACTGGCAGCGACTCATCAACAAC 291

Qy      102 AsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLysGlu 121
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      292 AACTGGGGCATCGACCCCAAGCCATGCGGGTCAAAATCTTCAACATCCAGGTCAAGGAG 351

Qy      122 ValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGlnVal 141
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      352 GTCACGACTCGAACCGCGAGACAACGGTGGCTAATAACCTTACCAGCACGGTTTCAGATC 411

Qy      142 PheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCysLeu 161
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      412 TTTCGGGACTCGTCGTACGAACACTGCCGTACGTGATGGATGGCGGTCAAGAGGGCAGCCTG 471

Qy      162 ProProPheProAlaAspValPheMetIleProGlnTyrGlyTyr-----LeuThr 178
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      472 CCTCCTTTTCCCAACGACGTCTTTATGGTGCCCCAGTACGGGCTACTGTGGACTGGTGACC 531
```

```
Qy      179 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 198
      |||
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      532 GGCAACACTTCGCAGCAACAGACTGACAGAAATGCTTCTACTGCTGGAGTACTTCTCT 591

Qy      199 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro 218
      |||
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      592 TCGCAGATGCTGCGGACTGGCAACAACCTTTGAAATTACGTACAGTTTGTAGAAGGTGCT 651

Qy      219 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 238
      |||
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      652 TTCCACTCGATGTACGCGCACAGCCAGAGCCTGGACCGGTGATGAACCCCTCTCATCGAC 711

Qy      239 GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsn-----GlnSerGlySerAla 254
      |||
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      712 CAGTACCTGTGGGACTGCAATCGACCACCACCGGAACCCCTGAATGCCGGGACTGCC 771

Qy      255 GlnAsnLysAspLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLys 274
      |||
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      772 ACCACCAAC-----TTTACCAAGCTGGCGCCTACCAACTTTTCCAACTTTAAAAAAG 822

Qy      275 AsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsn 294
      |||
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      823 AACTGGCTGCCCGGCTTCAATCAAGCAGCAGGGCTTCTCAAAGACTGCC-----AAT 876

Qy      295 AsnAsnSerAsnPheThrTrpThrGlyAlaSer-----LysTyr----- 307
      |||
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      877 CAAAACCTACAAGATCCCTGCCACCGGGTCAGACAGTCTCATCAATACGAGACGCACAGC 936

Qy      308 AsnLeuAsnGlyArgGluSerIleIleAsnProGlyThrAlaMetAlaSerHisLysAsp 327
      |||
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      937 ACTCTGGACGGAAGATGGAGTGCCTGACCCCGGACCTCCAATGGCCACGGCTGGACCT 996

Qy      328 AspGluAspLysPhePheProMetSerGlyValMetIlePheGlyLysGluSerAlaGly 347
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      997 GCGGACAGCAAG--TTCAGCAACAGCCAGCTCATCTTTGCGGGGCTAAACAGAACGSC 1053

Qy      348 AlaSerAsnThrAlaLeuAspAsnValMetIleThrAspGluGluGluIleLysAlaThr 367
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1054 AACACGGCCACCGTACCGGGACTCTGATCTTCACTCTGAGGAGGAGCTGGCAGCCACC 1113

Qy      368 AsnProValAlaThrGluArgPheGlyThrValAlaValAsnPheGlnSerSerSerThr 387
      |||
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1114 AACGCCACCGATACGGACATGTGGGGCAACCTACCTGCGGGTGACCCAGAGCAACAGCAAC 1173

Qy      388 AspProAlaThrGlyAspValHisAlaMetGlyAlaLeuProGlyMetValTrpGlnAsp 407
      |||
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1174 CTGCCGACCGTGGACAGACTGACAGCCTTGGGAGCGCTGGAATGGTCTGGCAAAAC 1233

Qy      408 ArgAspValTyrLeuGlnGlyProIleTrpAlaLysIleProHisThrAspGlyHisPhe 427
      |||
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1234 AGAGACATTTACTACCGGGTCCCATTTGGGCCAAGATTCTCATACCGATGGACACTTT 1293

Qy      428 HisProSerProLeuMetGlyGlyPheGlyLeuLysAsnProProGlnIleLeuIle 447
      |||
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1294 CACCCCTCACCGCTGATTGGTGGTTTGGGCTGAACACACCCGCTCTCTCAAATTTTATC 1353

Qy      448 LysAsnThrProValProAlaAsnProProAlaGluPheSerAlaThrLysPheAlaSer 467
      |||
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1354 AAGAACACCCCGGTACTCTGGAATCCTGCAACGACCTTCAGCTCTACTCCGGTAAACTCC 1413

Qy      468 PheIleThrGlnTyrSerThrGlyGlnValSerValGluIleGluTrpGluLeuGlnLys 487
      |||
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1414 TTCATTACTCAGTACAGCACTGGCCAGGTGTCCGGTGCAGATTGACTGGGAGATCCAGAAG 1473

Qy      488 GluAsnSerLysArgTrpAsnProGluValGlnTyrThrSerAsnTyrAlaLysSerAla 507
      |||
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1474 GAGCGGTCCAAACGCTGGAACCCCGAGGTCCAGTTTACCTCCAACTACGGACAGCAAAAC 1533

Qy      508 AsnValAspPheThrValAspAsnAsnGlyLeuTyrThrGluProArgProIleGlyThr 527
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      1534 TCTCTGTTGTGGGCTCCCGATGCGGTGGGAAATACACTGAGCCTAGGGCTATCGGTACC 1593

Qy      528 ArgTyrLeuThrArgProLeu 534
```



```
Db      1594 CGTACCTCACCCACCACCTG 1614
RESULT 11
US-09-532-594B-17
; Sequence 17, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.0252U2
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 1800
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: synthetic construct
; NAME/KEY: misc_feature
; LOCATION: 342
; OTHER INFORMATION: n = g, a, c or t(u)
; NAME/KEY: misc_feature
; OTHER INFORMATION: AAV4 capsid protein VP2 gene
US-09-532-594B-17

Alignment Scores:
Pred. No.:      5.02e-182      Length:      1800
Score:          1690.50      Matches:      325
Percent Similarity: 71.30%      Conservative: 65
Best Local Similarity: 59.41%      Mismatches:  132
Query Match:      58.17%      Indels:      25
DB:              4          Gaps:         9

US-09-807-802A-17 (1-534) x US-09-532-594B-17 (1-1800)

QY      2 AlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGlyAsn 21
Db      187 GCAGCAGCTGGCGGAGCTGCAGTC---GAGGSGGACAAAGTGGCGGATGGGTAAT 243
QY      22 AlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThrSer 41
Db      244 GCCTCGGGTGATTGGCAATTGGATTCCACCTGGTCTGAGGGGCCACGTCACGACCACG 303
QY      42 ThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSerAla 61
Db      304 ACCAGAACCTGGGTCTTGCCACCTACAACCAACCACTTACAAAGCGACTCGGAGAG--- 360
QY      62 SerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyrPhe 81
Db      361 -----AGCCTGCAGTCCAAACACCTACAACGGATTCTCCACCCCTGGGATACCTT 411
QY      82 AspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsnAsn 101
Db      412 GACTTCAACCGCTTCCACTGCCACTTCTCACCACTGCTGAGTGGCAGCGACTCATCAACAAC 471
QY      102 AsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLysGlu 121
Db      472 AACTGGGGCATCGGACCCCAAGCCATCGGGTCAAAATCTTCAACATCCAGGTCAAGGAG 531
QY      122 ValThrThrAsnAspGlyValThrIleAlaAsnAsnLeuThrSerThrValGlnVal 141
Db      532 GTACGACGTCGAACGGCGAGACAACGGTGGCTAATAACCTTACCAGCACGGTTCAGATC 591
QY      142 PheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCysLeu 161
Db      592 TTGCGGACTCGTCGTACGAACTGCGGTACGTGATGGATGCGGGTCAAGAGGGCAGCCTG 651
```

```
QY      162 ProProPheProAlaAspValPheMetIleProGlnTyrGlyTyr-----LeuThr 178
Db      652 CCTCCTTTTCCCAACGACGCTTTTATGGTGGCCCCAGTAGCGGCTACTGTGGACTGGTGACC 711
QY      179 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 198
Db      712 GGCACACACTTCGCAGCAACAGACTGACAGAAATGCTTCTACTGCTGGAGTACTTCTCT 771
QY      199 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro 218
Db      772 TCGCAGATGCTGCGGACTGGCAACAACTTTGAAATTACGTACAGTTTGTGAGAAAGTGCT 831
QY      219 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 238
Db      832 TTCCACTCGATGTACGCGCACAGCCAGAGCCTGGACCGCTGATGAACCCCTCTCATCGAC 891
QY      239 GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsn-----GlnSerGlySerAla 254
Db      892 CAGTACCTGTGGGACTGCAATCGACCACCACCGGAACACCCTGATGCGGGACTGCC 951
QY      255 GlnAsnLysAspLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLys 274
Db      952 ACCACCAAC-----TTTACCAAGCTGGCGCCTACCAACTTTTCCAACTTTTAAAAAG 1002
QY      275 AsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsn 294
Db      1003 AACTGGCTGCCCGGCTTCAATCAAGCAGCAGGGCTTCTCAAGACTGCC-----AAT 1056
QY      295 AsnAsnSerAsnPheThrTrpThrGlyAlaSer-----LysTyr----- 307
Db      1057 CAAAACACTACAAGATCCCTGCGCAGCGGTGACAGAGTCTCATCAATACGAGACGCACAGC 1116
QY      308 AsnLeuAsnGlyArgGluSerIleIleAsnProGlyThrAlaMetAlaSerHisLysAsp 327
Db      1117 ACTCTGGACGGAAGATGGAGTGGCTGACCTGACCCCGGACCTCCAATGGCCACGGTGGACCT 1176
QY      328 AspGluAspLysPhePheProMetSerGlyValMetIlePheGlyLysGluSerAlaGly 347
Db      1177 GCGGACAGCAAG---TTCAGCAACAGCCAGCTCATCTTTGCGGGGCTTAAACAGAACGGC 1233
QY      348 AlaSerAsnThrAlaLeuAspAsnValMetIleThrAspGluGluGluIleLysAlaThr 367
Db      1234 AACACGGCCACCGTACCCGGGACTCTGATCTTCACCTCTGAGGAGGAGCTGGCAGCCACC 1293
QY      368 AsnProValAlaThrGluArgPheGlyThrValAlaValAsnPheGlnSerSerThr 387
Db      1294 AACGCCACCGATACGGACATGTGGGGCAACCTACCTGGCGGTGACCAGAGCAACAGCAAC 1353
QY      388 AspProAlaThrGlyAspValHisAlaMetGlyAlaLeuProGlyMetValTrpGlnAsp 407
Db      1354 CTGCCGACCGTGGACAGACTGACAGCCTTGGGAGCCGTGCTTGGAAATGGTCTGGCAAAAC 1413
QY      408 ArgAspValTyrLeuGlnGlyProIleTrpAlaLysIleProHisThrAspGlyHisPhe 427
Db      1414 AGAGACATTTACTACAGGGTCCCATTTGGGCCAAGATTCTCATACCGATGGACACTTT 1473
QY      428 HisProSerProLeuMetGlyPheGlyLeuLysAsnProProGlnIleLeuIle 447
Db      1474 CACCCCTCACCGCTGATTGGTGGTGGTGGCTGAAACACCCCGCTCTCTCAAAATTTTATC 1533
QY      448 LysAsnThrProValProAlaAsnProProAlaGluPheSerAlaThrLysPheAlaSer 467
Db      1534 AAGAACACCCCGGTACTCTGCGAATCTCTGCAACGACCTTCAGCTCTACTCCGGTAAACTCC 1593
QY      468 PheIleThrGlnTyrSerThrGlyGlnValSerValGluIleGluTrpGluLeuGlnLys 487
Db      1594 TTCATTACTCAGTACAGCAGCTGGCCAGGTGTGGTGCAGATTGACTGGGAGATCCAGAAG 1653
QY      488 GluAsnSerLysArgTrpAsnProGluValGlnTyrThrSerAsnTyrAlaLysSerAla 507
Db      1654 GAGCGGTCCAAACGCTGGAACCCCGAGGTCCAGTTTACCTCCAACCTACGACAGCAAAAC 1713
```

Qy 508 AsnValAspPheThrValAspAsnAsnGlyLeuTyrThrGluProArgProIleGlyThr 527  
Db 1714 TCCTCTGTGGCTCCCGATGGGCTGGAAATACACTGAGCTAGGGCTATCGGTACC 1773  
Qy 528 ArgTyrLeuThrArgProLeu 534  
Db 1774 CGCTACTCACCCACCACCTG 1794  
RESULT 12  
US-09-532-594B-5  
; Sequence 5, Application US/09532594B  
; Patent No. 6468524  
; GENERAL INFORMATION:  
; APPLICANT: Chorini, John A.  
; APPLICANT: Kotin, Robert M.  
; APPLICANT: Safer, Brian  
; APPLICANT: Davidson, Beverly  
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF  
; FILE REFERENCE: 14014.0252U2  
; CURRENT APPLICATION NUMBER: US/09/532,594B  
; CURRENT FILING DATE: 2000-03-22  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 2208  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =  
; NAME/KEY: misc\_feature  
; LOCATION: 750  
; OTHER INFORMATION: n = g, a, c or t (u)  
; OTHER INFORMATION: AAV4 caspid protein VP1 gene  
US-09-532-594B-5  
Alignment Scores:  
Pred. No.: 7,07e-182 Length: 2208  
Score: 1690.50 Matches: 325  
Percent Similarity: 71.30% Conservative: 65  
Best Local Similarity: 59.41% Mismatches: 132  
Query Match: 58.17% Indels: 25  
DB: 4 Gaps: 9  
US-09-807-802A-17 (1-534) x US-09-532-594B-5 (1-2208)  
Qy 2 AlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGlyAsn 21  
Db 595 GCAGCAGCTGGCGAGCTGCAGTC--GAGGSGGACAAGGTGCCGATGGAGTGGTAAT 651  
Qy 22 AlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThrSer 41  
Db 652 GCCTCGGGTGATGGCATTGGGATTCACCTGGTCTGAGGGCCACGTACGACCACCGC 711  
Qy 42 ThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSerAla 61  
Db 712 ACCAGAACCTGGGTCTTGCCACCTACAAACACCACCTNTACAGCGACTCGGAGAG--- 768  
Qy 62 SerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyrPhe 81  
Db 769 -----AGCCTGCAGTCCAAACACCTACAACCGGATTCTCCACCCCTGGGATACCTT 819  
Qy 82 AspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsnAsn 101  
Db 820 GACTTCAACCGCTTCCACTGCCACTTCTCACCACGTGACTGGCAGCGACTCATCAACAAC 879  
Qy 102 AsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLysGlu 121  
Db 880 AACTGGGGCATGCCACCCAAAGCCATCGGGTCAAAATCTTCAACATCCAGGTCAAGGAG 939  
Qy 122 ValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGlnVal 141

Db 940 GTCACGACGTGGAACGGCGAGACAACGGTGGTAAATAACCTTACCAGCACGGTTCAGATC 999  
Qy 142 PheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCysLeu 161  
Db 1000 TTTGGGACTCGTCTAGAACTGCCGTACGTGATGGATGCGGGTCAAGAGGGCAGCCTG 1059  
Qy 162 ProProPheProAlaAspValPheMetIleProGlnTyrGlyTyr-----LeuThr 178  
Db 1060 CCTCCTTTTCCCAACGACGTCTTTATGGTGGCCAGTAGCGGCTACTGTGGACTGGTGACC 1119  
Qy 179 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 198  
Db 1120 GGCAACACTTCGCAGCAACAGACTGACAGAAATGCCTTCTACTGCCGTGGAGTACTTCTCT 1179  
Qy 199 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro 218  
Db 1180 TCGCAGATGCTGGGACTGGCAACACTTTGAAATATACGTACAGTTTGTGAGAAGGTGCCT 1239  
Qy 219 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 238  
Db 1240 TTCCACTCGATGTACGCGCACAGCCAGAGCCTGGACCGGCTGATGAACCTCTCATCGAC 1299  
Qy 239 GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsn-----GlnSerGlySerAla 254  
Db 1300 CAGTACCTGTGGGGACTGCAATCGACCACCACCGGAACCCCTGAATGCCGGGACTGCC 1359  
Qy 255 GlnAsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLys 274  
Db 1360 ACCACCAAC-----TTTACCAAGCTGGCGCTACCAACTTTTCCAACTTAAAAAG 1410  
Qy 275 AsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsn 294  
Db 1411 AACTGGCTGCCCGGCTTCAATCAAGCAGCAGGGGTTCTCAAAGACTGCC-----AAT 1464  
Qy 295 AsnAsnSerAsnPheThrTrpThrGlyAlaSer-----LysTyr----- 307  
Db 1465 CAAAACACTACAGATCCCTGCCACCGGGTCAGACAGTCTCATCAATACGAGACGCACAGC 1524  
Qy 308 AsnLeuAsnGlyArgGluSerIleIleAsnProGlyThrAlaMetAlaSerHisLysAsp 327  
Db 1525 ACTCTGGACGGAAGATGGAGTGCCTGACCCCGGACCTCCAATGGCCACGGCTGGACCT 1584  
Qy 328 AspGluAspLysPhePhePrometSerGlyValMetIlePheGlyLysGluSerAlaGly 347  
Db 1585 CGGACAGCAAG---TTCAGCAACAGCCAGCTCATCTTTGCGGGGCTAAACAGAACGGC 1641  
Qy 348 AlaSerAsnThrAlaLeuAspAsnValMetIleThrAspGluGluGluIleLysAlaThr 367  
Db 1642 AACACGGCCACCGTACCCGGGACTCTGATCTTCACTCTGAGGAGGAGCTGGCAGCCACC 1701  
Qy 368 AsnProValAlaThrGluArgPheGlyThrValAlaValAsnPheGlnSerSerThr 387  
Db 1702 AACGCCACCGATACGGACATGTGGGCAACCTACCTGGCGGTGACCCAGAGCAACAGCAAC 1761  
Qy 388 AspProAlaThrGlyAspValHisAlaMetGlyAlaLeuProGlyMetValTrpGlnAsp 407  
Db 1762 CTGCGACCGGTGGACAGACTGACAGCCTTGGGAGCCGCTGGAATGGTCTGGCAAAAC 1821  
Qy 408 ArgAspValTyrLeuGlnGlyProIleTyrAlaLysIleProHisThrAspGlyHisPhe 427  
Db 1822 AGAGACATTTACTACCAGGTCCCATTTGGGCCAAGATTCCTCATACCGATGGACACTTT 1881  
Qy 428 HisProSerProLeuMetGlyPheGlyLeuLysAsnProProGlnIleLeuIle 447  
Db 1882 CACCCCTCACCGCTGATTGGTGGGTTTGGGCTGAACACCCGCTCTCTCAAAATTTTATC 1941  
Qy 448 LysAsnThrProValProAlaAsnProProAlaGluPheSerAlaThrLysPheAlaSer 467  
Db 1942 AAGAACACCCCGGTACTCGCAATCCTGCAACGACCTTCAGCTCTACTCCGGTAAACTCC 2001  
Qy 468 PheIleThrGlnTyrSerThrGlyGlnValSerValGluIleGluTrpGluLeuGlnLys 487  
Db 2002 TTCATTACTCAGTACAGCACTGGCCAGGTGTCCGGTGCAGATTGACTGGGAGATCCAGAAG 2061





Db 4201 AAGAACACCCCGGTACCTGCGAATCCTGCAACGACCTTCAGCTCTACTCCGGTAAACTCC 4260

Qy 468 PheIleThrGlnTyrSerThrGlyGlnValSerValGluIleGluTrpGluLeuGlnLys 487

Db 4261 TTCAATTACTGATACAGCACTGGCCAGGTGTCGGTGCAAGATTGACTGGGAGATCCAGAAG 4320

Qy 488 GluAsnSerLysArgTrpAsnProGluValGlnTyrThrSerAsnTyrAlaLysSerAla 507

Db 4321 GAGCGGTCCAAACGCTGGAACCCCGAGGTCCAGTTTACCTCCAACACTACGGACAGCAAAAC 4380

Qy 508 AsnValAspPheThrValAspAsnAsnGlyLeuTyrThrGluProArgProIleGlyThr 527

Db 4381 TCTCTGTTGTGGGCTCCCGATGCGGCTGGGAAATACACTGAGCCTAGGGCTATCGGTACC 4440

Qy 528 ArgTyrLeuThrArgProLeu 534

Db 4441 CGCTACCTCACCCACCACCTG 4461

RESULT 14

US-09-438-268-1/c

; Sequence 1, Application US/09438268

; Patent No. 6491907

; GENERAL INFORMATION:

; APPLICANT: Rabinowitz, Joseph E.

; APPLICANT: Samulewski, Richard J

; APPLICANT: Xiao, Weidong

; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING

; TITLE OF INVENTION: THE SAME

; FILE REFERENCE: 5470-186

; CURRENT APPLICATION NUMBER: US/09/438,268

; CURRENT FILING DATE: 1999-11-10

; EARLIER APPLICATION NUMBER: 60/107,840

; EARLIER FILING DATE: 1998-11-10

; EARLIER APPLICATION NUMBER: 60/123,651

; EARLIER FILING DATE: 1999-03-10

; NUMBER OF SEQ ID NOS: 59

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 7214

; TYPE: DNA

; ORGANISM: Virus

US-09-438-268-1

Alignment Scores:

Pred. No.: 5.16e-181 Length: 7214

Score: 1690.50 Matches: 325

Percent Similarity: 71.30% Conservative: 65

Best Local Similarity: 59.41% Mismatches: 132

Query Match: 58.17% Indels: 25

DB: 4 Gaps: 9

US-09-807-802A-17 (1-534) x US-09-438-268-1 (1-7214)

Qy 2 AlaSerGlyGlyAlaProMetAlaAspAsnAsnGluGlyAlaAspGlyValGlyAsn 21

Db 2389 GCAGCAGCTGGCGAGCTGCAGTC---GAGGCGGACAAGGTGCCGATGGGTAAT 2333

Qy 22 AlaSerGlyAsnTrpHisCysAspSerThrTrpLeuGlyAspArgValIleThrThrSer 41

Db 2332 GCCTCGGGTGATTGGCATTGCCATTCCACCTGTGAGGGCCACGTCACGACCACCGC 2273

Qy 42 ThrArgThrTrpAlaLeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSerAla 61

Db 2272 ACCAGAACCTGGGTCTTGCCACCTACAAACACCACCTCTACAAAGCGACTCGGAGAG--- 2216

Qy 62 SerThrGlyAlaSerAsnAspAsnHisTyrPheGlyTyrSerThrProTrpGlyTyrPhe 81

Db 2215 -----AGCTGCAGTCCAAACACCTACACGGATTCTCCACCCCTGGGATACTTT 2165

Qy 82 AspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeuIleAsnAsn 101

Db 2164 GACTTCAACCGCTTCCACTGCCACTTCTTACCACGCTGACTGGCAGCGACTCATCAACAAC 2105

Qy 102 AsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGlnValLysGlu 121

Db 2104 AACTGGGCATCGCACCCCAAGCCATGCGGGTCAAAATCTTCAACATCCAGGTCAAGGAG 2045

Qy 122 ValThrThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSerThrValGlnVal 141

Db 2044 GTCACGACGTGCAACGGCGAGACAACGGTGGCTAATAACCTTACCAGCACGGTTCAGATC 1985

Qy 142 PheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHisGlnGlyCysLeu 161

Db 1984 TTTCGGGACTCGTCGTACGAACCTGCCGTACGTGATGGATGCGGGTCAAGAGGGCAGCCTG 1925

Qy 162 ProProPheProAlaAspValPheMetIleProGlnTyrGlyTyr-----LeuThr 178

Db 1924 CCTCCTTTTCCCAACGACGCTCTTTATGTTGCTGCCAGTACGGCTACTGTGGACTGGTGACC 1865

Qy 179 LeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhePro 198

Db 1864 GGCAACACTTCGCAGCAACAGACTGACAGAAATGCCTTCTACTGCTGGAGTACTTTCTCT 1805

Qy 199 SerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluValPro 218

Db 1804 TCGCAGATGCTCGGACTGCGCAACAACCTTTGAATATTACGTACAGTTTGTGAGAGGTGCT 1745

Qy 219 PheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIleAsp 238

Db 1744 TTCCACTCGATGTACGGCGACAGCCAGAGCCTGGACCGGCTGATGAACCTCTCTCATCGAC 1685

Qy 239 GlnTyrLeuTyrTyrLeuAsnArgThrGlnAsn-----GlnSerGlySerAla 254

Db 1684 CAGTACCTGTGGGACTGCAATCGACCACCACCACCCCTGAAATGCCGGGACTGCC 1625

Qy 255 GlnAsnLysAspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLys 274

Db 1624 ACCACCAAC-----TTTACCAAGCTGCGGCTACCAACTTTTCCAACTTTTAAAAAG 1574

Qy 275 AsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsn 294

Db 1573 AACTGGCTGCCCGGCTTCAATCAAGCAGCAGGGCTTCTCAAGAGACTGCC-----AAT 1520

Qy 295 AsnAsnSerAsnPheThrTrpThrGlyAlaSer-----LysTyr----- 307

Db 1519 CAAAACACTACAAGATCCCTGCCACCGGTCAGACAGTCTCATCAAAATACGAGACGCACAGC 1460

Qy 308 AsnLeuAsnGlyArgGluSerIleIleAsnProGlyThrAlaMetAlaSerHisLysAsp 327

Db 1459 ACTCTGGACGGAAGATGGAGTGGCTGACCCCGGACCTTCAATGGCCACGGCTGGACCT 1400

Qy 328 AspGluAspLysPhePheProMetSerGlyValMetIlePheGlyLysGluSerAlaGly 347

Db 1399 GCGGACAGCAAG--TTCAGCAACAGCCAGCTCATCTTTGCGGGGCTTAAACAGAACGGC 1343

Qy 348 AlaSerAsnThrAlaLeuAspAsnValMetIleThrAspGluGluIleLysAlaThr 367

Db 1342 AACACGGCCACCGTACCCGGGACTCTGTATCTTCACCTCTGAGGAGGAGCTGGCAGCCACC 1283

Qy 368 AsnProValAlaThrGluArgPheGlyThrValAlaValAsnPheGlnSerSerThr 387

Db 1282 AACGCCACCGATACGGACATGTGGGGCAACCTACCTGCGCGGTGACCAGAGCAACAGCAAC 1223

Qy 388 AspProAlaThrGlyAspValHisAlaMetGlyAlaLeuProGlyMetValTrpGlnAsp 407

Db 1222 CTGCGGACCGTGGACAGACTGACAGCCTTGGGAGCCCTGCGCTGGAATGGTCTGGCAAAAC 1163

Qy 408 ArgAspValTyrLeuGlnGlyProIleTrpAlaLysIleProHisThrAspGlyHisPhe 427

Db 1162 AGAGACATTTACTACCAAGGTCCCATTGGGCCAAAGATTCTCATATCCGATGGACACTTT 1103

Qy 428 HisProSerProLeuMetGlyGlyPheGlyLeuLysAsnProProGlnIleLeuIle 447

Db 1102 CACCCCTCACCGCTGATTGGTGGGTTTGGGCTGAAACACACCCGCTCTCTCAAAATTTTATC 1043

Qy 448 LysAsnThrProValProAlaAsnProProAlaGluPheSerAlaThrLysPheAlaSer 467

Db 1042 AAGAACACCCCGTACCTCGGAATCCTGCAACGACCTTCAGCTCTACTCCGGTAAACTCC 983  
Qy PheIleThrGlnTyrSerThrGlyGlnValSerValGluIleGluTrpGluLeuGlnLys 487  
Db 982 TTCATTACTCAGTACAGCACTGGCCAGGTGTCGGTGCAGATTGACTGGGAGATCCAGAAG 923  
Qy GluAsnSerLysArgTrpAsnProGluValGlnTyrThrSerAsnTyrAlaLysSerAla 507  
Db 922 GAGCGGTCCAAACGCTGGAACCCCGAGGTCCAGTTTACCTCCAACCTACGGACAGCAAAAC 863  
Qy 508 AsnValAspPheThrValAspAsnAsnGlyLeuTyrThrGluProArgProIleGlyThr 527  
Db 862 TCTCTGTGTGGCTCCCGATGGGCTGGGAAATACACTGAGCCTAGGCTATCGGTACC 803  
Qy 528 ArgTyrLeuThrArgProLeu 534  
Db 802 CGCTACCTCACCCACACCTG 782

RESULT 15

US-09-438-268-3  
; Sequence 3, Application US/09438268  
; Patent No. 6491907  
; GENERAL INFORMATION:  
; APPLICANT: Rabinowitz, Joseph E.  
; APPLICANT: Samulski, Richard J  
; APPLICANT: Xiao, Weidong  
; TITLE OF INVENTION: VIRUS VECTORS AND METHOD OF MAKING AND ADMINISTERING  
; TITLE OF INVENTION: THE SAME  
; FILE REFERENCE: 5470-186  
; CURRENT APPLICATION NUMBER: US/09/438,268  
; EARLIER FILING DATE: 1999-11-10  
; EARLIER APPLICATION NUMBER: 60/107,840  
; EARLIER FILING DATE: 1998-11-10  
; EARLIER APPLICATION NUMBER: 60/123,651  
; EARLIER FILING DATE: 1999-03-10  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 2271  
; TYPE: DNA  
; ORGANISM: Virus  
US-09-438-268-3

Alignment Scores:  
Pred. No.: 1.72e-45 Length: 2271  
Score: 491.50 Matches: 144  
Percent Similarity: 43.57% Conservative: 83  
Best Local Similarity: 27.64% Mismatches: 239  
Query Match: 16.91% Indels: 55  
DB: 4 Gaps: 16

US-09-807-802A-17 (1-534) x US-09-438-268-3 (1-2271)

Qy 10 AlaAspAsnAsnGluGlyAlaAspGlyValGlyAsnAlaSer-----GlyAsnTrpHis 27  
Db 625 GCAGAAGCCAGCACTGCTGCAGGAGGGGGGGCGAGTAATCTGTCAAAAGCATGTGGAGT 684  
Qy 28 CysAspSerThrTrpLeuGlyAspArgValIleThrThrSerThrArgThrTrpAlaLeu 47  
Db 685 GAGGGGCCACTTTTAGTGTAACTCTGTAACTTGTACATTTTCCAGACAGTTTAAATT 744  
Qy 48 ProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSerAlaSerThrGlyAlaSerAsn 67  
Db 745 CCATATGACCCAGAGCACCATTAATAAGGTGTTTCTCCGCAGCGAGTAGCTGCCACAAT 804  
Qy 68 Asp-----AsnHisTyrPheGlyTyrSerThrPro 77  
Db 805 GCCAGTGGAAAGGAGGCAAGGTTTGACCATCATGTCCTCCATAATGGGATACTCAACCCCA 864  
Qy 78 TrpGlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArg 97  
Db 865 TGCAGATATTAGATTTTAATGCTTTAAATTTATTTTTCACCTTTAGAGTTTCAGCAC 924

Qy 98 LeuIleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIle 117  
Db 925 TTAATTGAAAATTATGGAAGTATAGCTCCTGATGCTTTAACTGTAACTATATCAGAAATT 984  
Qy 118 GlnValLysGluValThr-----ThrAsnAspGlyValThrThrIleAlaAsnAsnLeu 135  
Db 985 GCTGTTAAGGATGTTACAGACAAAACCTGGAGGGGGTA---CAGGTTACTGACGACACT 1041  
Qy 136 ThrSerThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySer 155  
Db 1042 ACAGGGCGCCTATGCATGTTAGTACACCATGAATACAAAGTACCCATATGTGTTAGGCA 1101  
Qy 156 AlaHisGlnGlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGly 175  
Db 1102 GGTACGATACTTTAGCCCCCAGAACTTCCTATTTGGGTATACTTTCCCTCAATATGCT 1161  
Qy 176 TyrLeuThrLeuAsnAsn---GlySerGlnAlaValGly----- 187  
Db 1162 TACTTAACAGTAGGAGATGTTAAACACACAAGGAATTTCTGGAGACAGCAAAATTAGCA 1221  
Qy 188 -----ArgSerSerPheTyrCysLeuGluTyrPheProSerGlnMetLeuArgThrGly 205  
Db 1222 AGTGAAGAATCAGCATTTTATGTTTTGGAACACAGTTCTTTTCAGCTTTTAGGTACAGA 1281  
Qy 206 AsnAsnPheThrPheSerTyrThrPheGluGluValProPheHisSerSerTyrAlaHis 225  
Db 1282 GGTACAGCAACTATGTCTTATAAGTTTCTCCAGTGCCTCCAGAAAATTTAGAGGGCTGC 1341  
Qy 226 SerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyrLeuTyrTyrLeuAsn 245  
Db 1342 AGTCAACACTTTTATGAAATGTACAATCCCTTA-----TACGGATCCCGCTTAGGG 1392  
Qy 246 ArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeuPheSerArgGlySer 265  
Db 1393 GTTCTCTGACACATTAGGAGGTGACCCAAATTTAGATCTTTA-----ACA 1437  
Qy 266 ProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyProCysTyrArgGlnGln 285  
Db 1438 CATGAAGACCATGCAATTTCAGCCCCCAAACTTCATGCCAGGSCCACTAGTAAACTCAGTG 1497  
Qy 286 ArgValSerLysThrLysThrAspAsnAsnAsnSerAsnPheThrTrpThrGlyAlaSer 305  
Db 1498 TCTACAAAGGAGGAGACAGCTCTAATACTGGAGCTGGAAAAGCCTTAACAGGCCCTTAGC 1557  
Qy 306 LysTyrAsnLeuAsnGlyArgGluSerIleIleAsnProGly---ThrAlaMetAlaSer 324  
Db 1558 ACAGGTACCTCTCAAAACACTAGATAATCTCTACGCCCTGGGCCAGTGTCTCAGCCATAC 1617  
Qy 325 HisLysAspAspGluAspLysPhePheProMetSerGlyValMetIlePheGlyLysGlu 344  
Db 1618 CACCACTGGGACACAGATAAATATGTACAGGAATAAATGCCATTTCTCATGGTTCAGACC 1677  
Qy 345 SerAlaGlyAlaSerAsnThrAlaLeuAspAsnVal-----Met 357  
Db 1678 ACTTATGCT-----AACGCTGAAGACAAAGAGTATCAGCAAGGAGTGGGTAGATTT 1728  
Qy 358 IleThrAspGluGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThr 377  
Db 1729 CCAATGAAAAAGAACAGCTAAACAGTTACAGGGTTTAAACATGCACACCTAC----- 1782  
Qy 378 ValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMet 397  
Db 1783 -----TTTCCCAATAAAGGAACCCAGCAATATACAGATCAAAATTGAG---CGC 1827  
Qy 398 GlyAlaLeuProGlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrp 417  
Db 1828 CCCCTAATGGTGGTCTGTATGGAACAGAAAGAGCCCTTCACTATGAAAGCCAGCTGTGG 1887  
Qy 418 AlaLysIleProHisThrAspGlyHisPheHisProSer---ProLeuMetGlyGlyPhe 436  
Db 1888 AGTAAATTCAAAATTTAGATGACAGTTTAAACTTAAAACTCAGTTTGCAGCCTTAGGAGGATGG 1947





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 14, 2004, 15:26:50 ; Search time 2963.43 Seconds  
(without alignments)  
4379.586 Million cell updates/sec

Title: US-09-807-802A-17  
Perfect score: 2906  
Sequence: 1 MASGGGAPMADNNEGADGVG.....NNGLYTEPRPIGTRYLTRPL 534

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09807802/runat\_11022004\_175608\_15941/app\_query.fasta\_1.2389  
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09807802@cgn\_1\_1\_7257@runat\_11022004\_175608\_15941 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID       | Description        |
|------------|-------|-------|--------|----|----------|--------------------|
| 1          | 134   | 4.6   | 856    | 28 | BH164736 | BH164736 ENTTS28TF |
| 2          | 125   | 4.3   | 754    | 10 | BG207226 | BG207226 RST26694  |
| 3          | 121.5 | 4.2   | 925    | 28 | BH152154 | BH152154 ENTPV12TR |
| 4          | 112   | 3.9   | 753    | 28 | BH115587 | BH115587 RPCI-24-3 |
| 5          | 110.5 | 3.8   | 768    | 14 | CD558053 | CD558053 AGENCOURT |
| 6          | 110.5 | 3.8   | 815    | 10 | BG746848 | BG746848 602704187 |
| 7          | 110.5 | 3.8   | 5809   | 11 | AK076994 | AK076994 Mus muscu |
| 8          | 109   | 3.8   | 687    | 13 | BU723756 | BU723756 SJMAZH04  |
| 9          | 106.5 | 3.7   | 754    | 10 | BE908188 | BE908188 601500441 |
| 10         | 106.5 | 3.7   | 786    | 10 | BE733520 | BE733520 601566055 |
| 11         | 106.5 | 3.7   | 830    | 13 | BQ888825 | BQ888825 AGENCOURT |
| 12         | 106.5 | 3.7   | 986    | 13 | BQ682489 | BQ682489 AGENCOURT |
| 13         | 105   | 3.6   | 606    | 13 | BU800802 | BU800802 SJF2CVC08 |
| 14         | 103.5 | 3.6   | 775    | 12 | BI253826 | BI253826 602976195 |
| 15         | 103.5 | 3.6   | 865    | 10 | BF340815 | BF340815 602037709 |
| 16         | 103   | 3.5   | 906    | 13 | BU848986 | BU848986 AGENCOURT |
| 17         | 102.5 | 3.5   | 626    | 13 | BQ416589 | BQ416589 IK39a02.x |
| 18         | 102.5 | 3.5   | 2742   | 11 | AK004837 | AK004837 Mus muscu |
| 19         | 102   | 3.5   | 829    | 13 | BU365695 | BU365695 603790568 |
| 20         | 102   | 3.5   | 2637   | 11 | BC037597 | BC037597 Mus muscu |
| 21         | 102   | 3.5   | 2688   | 11 | BC019128 | BC019128 Mus muscu |
| 22         | 101.5 | 3.5   | 645    | 12 | BM771938 | BM771938 K-EST0055 |
| 23         | 101   | 3.5   | 838    | 28 | AZ686611 | AZ686611 ENTLM95TF |
| 24         | 101   | 3.5   | 3057   | 11 | AK044947 | AK044947 Mus muscu |
| 25         | 100   | 3.4   | 688    | 12 | BM166442 | BM166442 EST568965 |
| 26         | 100   | 3.4   | 959    | 14 | CD458291 | CD458291 FG08_09b0 |
| 27         | 100   | 3.4   | 962    | 13 | BQ684616 | BQ684616 AGENCOURT |
| 28         | 100   | 3.4   | 1201   | 13 | BF128255 | BF128255 BX442505  |
| 29         | 100   | 3.4   | 2036   | 10 | BF128255 | BF128255 BX442505  |
| 30         | 100   | 3.4   | 4479   | 11 | BC044003 | BC044003 Xenopus 1 |
| 31         | 99.5  | 3.4   | 672    | 29 | CC198495 | CC198495 ZMMBCC031 |
| 32         | 99.5  | 3.4   | 691    | 29 | BZ315618 | BZ315618 Ia61g03.b |
| 33         | 99.5  | 3.4   | 739    | 14 | CA498545 | CA498545 WHE3244 G |
| 34         | 99.5  | 3.4   | 852    | 14 | CB990709 | CB990709 AGENCOURT |
| 35         | 99.5  | 3.4   | 894    | 12 | BI759543 | BI759543 603046955 |
| 36         | 99.5  | 3.4   | 950    | 13 | BU172290 | BU172290 AGENCOURT |
| 37         | 99    | 3.4   | 871    | 10 | BF269841 | BF269841 GA_EB000  |
| 38         | 99    | 3.4   | 912    | 13 | BX329591 | BX329591 BX329591  |
| 39         | 99    | 3.4   | 1039   | 29 | CNS077AT | AL432459 T3 end of |
| 40         | 99    | 3.4   | 2029   | 11 | AY104007 | AY104007 Zea mays  |
| 41         | 99    | 3.4   | 3329   | 11 | AK033012 | AK033012 Mus muscu |
| 42         | 98.5  | 3.4   | 841    | 13 | BX392614 | BX392614 BX392614  |
| 43         | 98.5  | 3.4   | 932    | 13 | BU838941 | BU838941 AGENCOURT |
| 44         | 98.5  | 3.4   | 1201   | 9  | AL520051 | AL520051 AL520051  |
| 45         | 98    | 3.4   | 704    | 9  | AL043249 | AL043249 DKFZp434K |

ALIGNMENTS

RESULT 1  
BH164736  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

BH164736  
ENTTS28TF Entamoeba histolytica Sheared DNA  
genomic, genomic survey sequence.

BH164736.1 GI:15738174  
GSS.  
Entamoeba histolytica  
Entamoeba histolytica  
Eukaryota; Entamoebidae; Entamoeba.  
1 (bases 1 to 856)  
Loftus,B., Wang,Z., Van Aken,S. and Fraser,C.

856 bp  
DNA  
linear  
GSS 24-SEP-2001







|  |     |   |     |
|--|-----|---|-----|
| Best Local Similarity: 20.14%                |     | Mismatches: 134   |     |
| Query Match: 4.18%                           |     | Indels: 154   |     |
| DB: 28                                       |     | Gaps: 21  |     |
| US-09-807-802A-17 (1-534) x BH152154 (1-925) |     |   |     |
| Qy   | 47  | LeuProThrTyrAsnAsnHisLeuTyrLysGlnIleSerSerAlaSerThrGlyAlaSer  | 66  |
| Db   | 29  | CTATCATCTCTCGATGAACAT-----CAATTCTCTTCATTTCAATCTGGTCAACAA      | 79  |
| Qy   | 67  | Asn-----AspAsnHisTyrPheGlyTyrSerThrProTrp                     | 78  |
| Db   | 80  | AATCCTTTTGGAACTCCAAACAATACTACACAACCTCAATTGGTACAAATTTCTCATTC   | 139 |
| Qy   | 79  | GlyTyrPheAspPheAsnArgPheHisCysHisPheSerProArgAspTrpGlnArgLeu  | 98  |
| Db   | 140 | GCTGGATTTAATACTAACCACAAACGTCGGGAACAACCTAATCCG-----            | 181 |
| Qy   | 99  | IleAsnAsnAsnTrpGlyPheArgProLysArgLeuAsnPheLysLeuPheAsnIleGln  | 118 |
| Db   | 182 | -----TTCAATTCTACT   | 193 |
| Qy   | 119 | VallysGluValThr---ThrAsnAspGlyValThrThrIleAlaAsnAsnLeuThrSer  | 137 |
| Db   | 194 | GGAAGTTCAGTCTCAGGAGCAATAAACCTTTTGTCTACTACTCAAAACAATAACAACAACT | 253 |
| Qy   | 138 | ThrValGlnValPheSerAspSerGluTyrGlnLeuProTyrValLeuGlySerAlaHis  | 157 |
| Db   | 254 | -----CCTTTCACAACAGGAACAACCTAAC                                | 277 |
| Qy   | 158 | GlnGlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGlyTyrLeu     | 177 |
| Db   | 278 | -----CCATTTAAT  | 286 |
| Qy   | 178 | ThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyrPhe  | 197 |
| Db   | 287 | ACTACTAACAATACTACTACTCA-----                                  | 310 |
| Qy   | 198 | ProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGluVal  | 217 |
| Db   | 311 | -----ACAACCTAATCCATTTAATACTACTAATACTACAAGCACCAACAAT           | 352 |
| Qy   | 218 | ProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeuIle  | 237 |
| Db   | 353 | CCATTTAATAACAATACTACTAAT-----AATCCATTT---                     | 385 |
| Qy   | 238 | AspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLys  | 257 |
| Db   | 386 | -----AGTACTAATAACACTACCTACCAATGGAAACAACCAAT---                | 418 |
| Qy   | 258 | AspLeuLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeu  | 277 |
| Db   | 419 | -----TTATTAAATCAA---ACTCCTTCTTCTAATACTATAACACCAGGAAT---       | 463 |
| Qy   | 278 | ProGlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnSer     | 297 |
| Db   | 464 | -----AATACAACCTACTACAACAGGAAGTAATCCATTTGGT                    | 499 |
| Qy   | 298 | AsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIleIleAsn  | 317 |
| Db   | 500 | AATTTCACTACCCTAAT-----AATACTACCTCCAGTACAACCTTCA               | 541 |
| Qy   | 318 | ProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMetSerGly  | 337 |
| Db   | 542 | ACAGGAACAACCATCACACAGGAAGTAATCCATTTGGTAATTTCACTCCCTCAAACTCA   | 601 |
| Qy   | 338 | ValMetIlePheGlyLysGluSerAlaGlyAlaSerAsnThrAlaLeuAspAsnValMet  | 357 |
| Db   | 602 | GCTCCAACAACAGGAATAATAACAACAGGAACAACCTACCACAACAGGAATAATACTAACT | 661 |
| Qy   | 358 | IleThrAspGluGluIleLysAlaThrAsnProValAlaThrGluArgPheGlyThr     | 377 |
| Db   | 662 | TCAACA-----GGAAGTAATCCA-----TTTGGT---                         | 685 |

|   |     |  |     |  |  |
|---|-----|--|-----|--|--|
| Qy  | 378 | ValAlaValAsnPheGlnSerSerSerThrAspProAlaThrGlyAspValHisAlaMet | 397 |  |  |
| Db  | 686 | -----AATTCACTACTTCAAACCTCACTACCACAGGAACGACAGCTACACA          | 736 |  |  |
| Qy  | 398 | GlyAlaLeuProGlyMetValTrpGlnAspArgAsp-ValTyrLeuGlnGlyProIleTr | 417 |  |  |
| Db  | 737 | GGA-----AATGACACAACCTACAGGACCA-----                          | 761 |  |  |
| Qy  | 417 | pAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGl | 437 |  |  |
| Db  | 762 | -----CCCCACACACAGGGAGTATCCAT-----TTGGGTAATTTTC--             | 797 |  |  |
| Qy  | 437 | YLeuLysAsnProProGlnIleLeuIleLysAsnThrProValProAlaAsnProPr    | 457 |  |  |
| Db  | 798 | -ATACAAATAATACCACTTCTATGCTAGTACAACTCCCAACACAGGAAGTAACCCACT   | 856 |  |  |
| Qy  | 457 | oAlaGluPheSerAlaThr  | 463 |  |  |
| Db  | 857 | TGGGTATTTCACCTACTCTCA  | 875 |  |  |
| RESULT 4  |     |  |     |  |  |
| BH115587  |     |  |     |  |  |
| LOCUS   |     |  |     |  |  |
| DEFINITION  |     |  |     |  |  |
| ACCESSION   |     |  |     |  |  |
| VERSION   |     |  |     |  |  |
| KEYWORDS  |     |  |     |  |  |
| SOURCE  |     |  |     |  |  |
| ORGANISM  |     |  |     |  |  |
| REFERENCE   |     |  |     |  |  |
| AUTHORS   |     |  |     |  |  |
| TITLE   |     |  |     |  |  |
| JOURNAL   |     |  |     |  |  |
| COMMENT   |     |  |     |  |  |
| FEATURES  |     |  |     |  |  |
| source  |     |  |     |  |  |
| 1..753  |     |  |     |  |  |
| /organism="Mus musculus"  |     |  |     |  |  |
| /mol_type="genomic DNA"   |     |  |     |  |  |
| /strain="C57BL/6J"  |     |  |     |  |  |
| /db_xref="taxon:10090"  |     |  |     |  |  |
| /clone="RPCI-24-358F16"   |     |  |     |  |  |
| /sex="Male"   |     |  |     |  |  |
| /cell_type="Spleen/Brain"   |     |  |     |  |  |
| /clone_lib="RPCI-24"  |     |  |     |  |  |
| /note="Vector: pTARBAC1; Site 1: BamH1; Site 2: BamH1; RPCI-24 Mouse BAC Library produced by Pieter de Jong. The library was cloned in the pTARBAC1 cloning vector at the BamH1 sites using MboI partially digested male C57BL/6J DNA." |     |  |     |  |  |
| BASE COUNT  |     |  |     |  |  |
| ORIGIN  |     |  |     |  |  |
| 211 a 178 c 167 g 197 t   |     |  |     |  |  |

BH115587 753 bp DNA linear GSS 19-JUL-2001  
RPCI-24-358F16.TV RPCI-24 Mus musculus genomic clone RPCI-24-358F16  
, genomic survey sequence.  
BH115587  
BH115587.1 GI:14954954  
GSS.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 753)  
Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,  
Tsagaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,  
Russell,D., de Jong,P. and Fraser,C.M.  
Mouse BAC End Sequences from Library RPCI-24  
Unpublished  
Other\_GSSs: RPCI-24-358F16.TJ  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org  
Clones are derived from the mouse BAC library RPCI-24. For BAC  
library availability, please contact Pieter de Jong  
(pdejong@mail.cho.org). Clones may be purchased from BACPAC  
Resources (<http://www.chori.org/bacpac/orderingframe.htm>). BAC end  
page: [http://www.tigr.org/tdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html)  
Plate: 358 row: F column: 16  
Seq primer: T7  
Class: BAC ends.

Location/Qualifiers  
1..753  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCI-24-358F16"  
/sex="Male"  
/cell\_type="Spleen/Brain"  
/clone\_lib="RPCI-24"  
/note="Vector: pTARBAC1; Site 1: BamH1; Site 2: BamH1; RPCI-24 Mouse BAC Library produced by Pieter de Jong. The library was cloned in the pTARBAC1 cloning vector at the BamH1 sites using MboI partially digested male C57BL/6J DNA."



cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: NIH Intramural Sequencing Center  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLCM1709 row: n column: 21  
High quality sequence stop: 809.

FEATURES

source

Location/Qualifiers  
1. .815  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4857428"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_15"  
/note="Organ: colon; Vector: pOTB7; Site 1: XhoI; Site 2:  
ECORI; cDNA made by oligo-dT priming. Directionally  
cloned into EcorI/XhoI sites using the following 5'  
adaptor: GGCACGAG(G). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)"  
176 a 246 c 243 g 149 t 1 others

BASE COUNT  
ORIGIN

Alignment Scores:

Pred. No.: 0.998 Length: 815  
Score: 110.50 Matches: 69  
Percent Similarity: 36.74% Conservative: 46  
Best Local Similarity: 22.04% Mismatches: 124  
Query Match: 3.80% Indels: 74  
DB: 10 Gaps: 11

US-09-807-802A-17 (1-534) x BG746848 (1-815)

Qy 197 PheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGlu 216  
Db 802 TTCGGCAACGAGGAGAGAACTAGTCTGAGTTTCTTTGTAGTGTGTCTTAGAA 743  
Qy 217 ValProPheHisSerTyrAlaHisSerGlnSerLeu-----AspArgLeuMet 233  
Db 742 ATAGGACATTATTCGGTTTAAACAATTAATAAATCTGTCTGGCAGGCCCTTTGCTC 683  
Qy 234 AsnProLeuIleAspGlnTyrLeuTyrTyrLeuAsnArgThrGln----- 248  
Db 682 AGGCCTGTGTCTCT-----GCTCAGCCACTCTCTGCACCACACTGACCCAGGGGCTGC 629  
Qy 249 -----AsnGlnSerGlySerAlaGlnAsnLysAspLeuLeuPheSerArgGlySerPro 266  
Db 628 CCGGACATAGAACGGAGGGGTTTCCGGGCCCCACTCCCTGCATGGCCGACGCCACCC 569  
Qy 267 AlaGlyMetSerValGlnProLysAsnTrpLeuProGlyProCysTyrArgGlnArg 286  
Db 568 GGGCTGTGCCACTACAGCCGGCTCACTGGCTCCAGGGACCACGCTCCACCCATACAG 509  
Qy 287 ValSerLysThrLysThrAspAsnAsnSerAsnPheThrTrpThrGlyAlaSerLys 306  
Db 508 CTTTCATCTGTGCCAGGTCCCTCTGGTCAAAAGCTCTGTGTGTATGGCCAGGG----- 458  
Qy 307 TyrAsnLeuAsnGlyArgGluSerIleIleAsnProGlyThrAlaMetAlaSerHis-Ly 326  
Db 457 -----CCTGGCACAGCTTGGAGGGGCCACTGC 431  
Qy 326 sAspAspGluAspLysPhePheProMetSerGlyValMetIlePheGlyLysGluSerAl 346  
Db 430 AGAG-----CTCGCGGTCTCTGAGGACACGGC 404  
Qy 346 aGlyAlaSerAsnThrAlaLeuAspAsnValMetIleThrAspGluGluIleLysAl 366  
Db 403 TGGCGGTGCC-----TTTCCGGAGGGTGTCTGCGAAGCT 371

Qy 366 aThrAsnProValAlaThrGluArgPheGlyThrValAlaValAsnPheGlnSerSerSe 386  
Db 370 GAGGCATGGTCTCCAGACC-----TTCCAGGGGCTCCAGTCTCGCAGCAAGCAAG 317  
Qy 386 rThrAspProAlaThrGlyAspValHisAlaMetGlyAlaLeuProGlyMetValTrpGl 406  
Db 316 CCCCCTCCCCCTGGCTGGAGATGTTTCATGCA-----GAAGTACATGC 275  
Qy 406 nAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIleProHisThrAspGlyHi 426  
Db 274 CGTAAATGATGTACACGTACAGGTCCCGGCTTCCATGAACATGCCTCGGTTGCGGGGG 215  
Qy 426 sPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsnProProProGlnIleLe 446  
Db 214 TCTGCCGGCCACCCCTTGAGTGGCGGCTTTCATCTCTCTGG-----CCCCAGGTATG 164  
Qy 446 uIleLysAsnThrProValProAlaAsnProProAlaGluPheSerAlaThrLysPheAl 466  
Db 163 CCTC-----GGTCTCCACGATCGGCCTCGGAGTTCTG 131  
Qy 466 aSerPhe-----IleThrGlnTyrSerThrGlyGlnValSerVa 479  
Db 130 TGCCATTAGGAAGTCGCGGACTAGGACCTGTCCAGAAATGCCGGGCCAGGGGGACTG 71  
Qy 479 lGluIleGluTrpGluLeuGlnLysGluAsnSerLys 491  
Db 70 CCGCTGGTGAAGAACTCCAAACCCAGTCGGGTAAG 34

RESULT 7  
AK076994

LOCUS  
DEFINITION

AK076994 5809 bp mRNA linear HTC 07-DEC-2002  
Mus musculus adult male testis cDNA, RIKEN full-length enriched  
library, clone:4931406P16 product:hypothetical protein, full insert  
sequence.

ACCESSION  
AK076994

VERSION  
KEYWORDS

AK076994.1 GI:26345807  
HTC; CAP trapper.

SOURCE  
ORGANISM

Mus musculus (house mouse)

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

1 Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
10349636  
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159  
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861  
4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,  
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,





```

Db      2749 AGCTTGGTCTAGTGAGCAGCTATATGGACAATATGATGTCAGAAAGTTTGGGACAGAAG 2808
Qy      157 HisGlnGlyCysLeuProPheProAlaAspValPheMetIleProGlnTyrGlyTyr 176
Db      2809 CCCAGGGA-----CCTAGAAATAAC--- 2829
Qy      177 LeuThrLeuAsnAsnGlySerGlnAlaValGlyArgSerSerPheTyrCysLeuGluTyr 196
Db      2830 ---ACCTGGCCAAACCGTGACCAAGTGATGGA----- 2859
Qy      197 PheProSerGlnMetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGluGlu 216
Db      2860 -----GTCCTTGGAAATGCTGGGAGAGATT 2883
Qy      217 ValProPheHisSerSerTyrAlaHisSerGlnSerLeuAspArgLeuMetAsnProLeu 236
Db      2884 TTGCCTTTTGATCCTGCAGTGGGC-----TCAGACCCAGAA 2919
Qy      237 IleAspGlnTyrLeuTyrTyrLeuAsnArgThrGlnAsnGlnSerGlySerAlaGlnAsn 256
Db      2920 TTGCACGCTATGTGGCAGGAGTGAGCCAGGCAATGCAGCAGAGGCGGCAAGCTCAACAT 2979
Qy      257 LysAspLeuPheSerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrp 276
Db      2980 GGCCGC-----CGGCCAGGCAACCCCGTGGC-----AACTGG 3012
Qy      277 LeuPro---GlyProCysTyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsn 295
Db      3013 CGGCCCATGGAAGACGCTCATCGGACCTTGGCCCTTGGCAGAGTCTTCACAGAAGGGGAC 3072
Qy      296 AsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsnLeuAsnGlyArgGluSerIle 315
Db      3073 AGCTTGCACAGCGGCTGGTCAGGAGCT-----CAGGAGAGACTCGGCCAGT 3117
Qy      316 IleAsnProGlyThrAlaMetAlaSerHisLysAspAspGluAspLysPhePheProMet 335
Db      3118 TCCAGTGATGAGACCTCTTCAGCC-----AATGGGACAGACTGTGTTCTCCATG 3165
Qy      336 -----SerGlyValMetIlePheGlyLysGluSerAlaGlyAla 348
Db      3166 TTTTCAGGGCCTGACCTTGTTGCTGCTGTATAACAGAGAGGAACACAGCAGTGGAGAG 3225
Qy      349 SerAsnThrAlaLeu-----AspAsnValMetIleThrAspGluGluGluIleLysAla 366
Db      3226 CAGGAGACCAGCACACTGCCCTCACCACCTCTCCTGACCACAGTGGAGGACGTGAACCAG 3285
Qy      367 ThrAsnProValAlaThrGluArgPheGlyThrValAlaValAsnPheGlnSerSer 386
Db      3286 GATAACAAACCAAAACG-----TGGCCACCCAAAGCACCATGGCAGCACCCCTTCC 3336
Qy      387 ThrAspProAlaThr-----GlyAspValHisAlaMetGlyAlaLeuPro 401
Db      3337 CCGATGCCAGCAGCGCTGCCAGCCAGCTGCACCACCTCTATGCAGTC---GCCAGTCCT 3393
Qy      402 GlyMetValTrpGlnAspArgAspValTyrLeuGlnGlyProIleTrpAlaLysIlePro 421
Db      3394 GGCAGCCAGTGGAAACGACACCCGTCAGATGCTGCAGTCCCCTGTGTGGGCTGCAGCCAAAT 3453
Qy      422 HisThrAspGlyHisPheHisProSerProLeuMetGlyGlyPheGlyLeuLysAsn-Pr 441
Db      3454 GACTGC-----AATACC 3465
Qy      441 OProProGlnIleLeuIleLysAsnThrProValProAlaAsnProProAlaGluPheSe 461
Db      3466 ACCTCCTTCACC-TATGTTTCAGACACCACCCAGCCCCCACCCTCCACCAGCACATAAGGC 3524
Qy      461 rAlaThrLys---PheAlaSerPhe 468
Db      3525 AGCACCCAAAGGGCTTCAAGGCCTTC 3549

```

```

RESULT 8
BU723756          687 bp      mRNA      linear      EST 12-FEB-2003
LOCUS

```

```

DEFINITION  SJMAZH04 SJM Schistosoma japonicum cDNA, mRNA sequence.
ACCESSION   BU723756
VERSION     BU723756.1  GI:28331125
KEYWORDS    EST.
SOURCE      Schistosoma japonicum
ORGANISM    Schistosoma japonicum
REFERENCE   1 (bases 1 to 687)
AUTHORS     Hu,W., Yan,Q., Shen,D., Liu,F., Xu,X., Zhu,Z., Zhang,X., Wang,J.,
Sun,J., Xu,X., Wang,Z., Zeng,L., Rong,Y., Wu,X., Qu,J., Xu,Z.,
Huang,J., Ma,Y., Wang,S., Wang,Z., Xue,C., Feng,Z., Chen,Z. and Han
,Z.
TITLE       Expressed sequence tags from male adults of Schistosoma japonicum
JOURNAL     Unpublished
COMMENT     Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn.
FEATURES             Location/Qualifiers
     source           1..687
                     /organism="Schistosoma japonicum"
                     /mol_type="mRNA"
                     /db_xref="taxon:6182"
                     /sex="male"
                     /tissue_type="Whole body"
                     /dev_stage="adult"
                     /lab_host="rabbits"
                     /clone_lib="SJM"
BASE COUNT  266 a   103 c   56 g   262 t
ORIGIN

```

```

Alignment Scores:
Pred. No.:      1.1      Length:      687
Score:          109.00    Matches:     50
Percent Similarity: 36.05% Conservative: 34
Best Local Similarity: 21.46% Mismatches:    63
Query Match:     3.75%   Indels:      86
DB:              13      Gaps:       13

```

US-09-807-802A-17 (1-534) x BU723756 (1-687)

```

Qy      111 AsnPheLysLeuPheAsnIleGlnValLysGluValThrThrAsnAspGlyValThrThr 130
Db      27  AACTATAATATTCTCAATTCTTCTTTA---CATACATCATCGTCATCATCAATAACA 83
Qy      131 IleAlaAsnAsnLeuThrSerThrValGlnValPheSerAspSerGluTyrGlnLeu--- 149
Db      84 ATAGTACCATCATTTATCATCT-----TATTCTCATTCACCGTTTAGTTTAAAT 131
Qy      150 -----ProTyrValLeuGlySerAlaHisGlnGlyCysLeuProPheProAlaAsp 167
Db      132 CCATTTCATATATATTAAATTTTCAAAAAAATCT-----TCAGAT 173
Qy      168 ValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsnAsnGlySerGlnAlaVal--- 186
Db      174 TTTTATTATTAAACAAAAAATCAAATTTTCTACAAATTTGGTTTCATATCATTTACCA 233
Qy      187 -----GlyArgSerSerPheTyrCysLeuGlu----- 195
Db      234 GAATTAATACGTACATTACAAGTCGTAAATCATTTATGGTGTTCACAAAATTGAATTTA 293
Qy      196 -----TyrPhePro-----SerGln 200
Db      294 TTTCACCATTTTTCATTTTATTTTCCATTAATTTATGATTTACAATTTTCTACACAT 353
Qy      201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu----- 215
Db      354 TTATTATCACTAATTAATAAATTTGAAATTTTCTACACAACTTAAATGATTTTATGTGAT 413

```

QY 216 -----GluValProPheHisSerSerTyrAlaHis 225  
Db 414 TTACCAGAAATTGAATATTGAATCGTGCACGTATACCATATCATCGT----- 461  
QY 226 SerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyrLeu---TyrTyrLeu 244  
Db 462 -----CAACAATATGTAAACATATATGTT 485  
QY 245 AsnArgThrGlnAsnGlnSerGlySerAlaGlnAsnLysAspLeuLeuPheSerArgGly 264  
Db 486 AATAATAATAATAATGATAAT-----ATTGCTAAAGAATTACGTTTTTTCACAT--- 533  
QY 265 SerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyProCysTyrArgGln 284  
Db 534 CCTCAAGCAAAACATTTCGTAAAT-----TATCAACAC 566  
QY 285 GlnArgValSerLysThrLysThrAspAsnAsnAsnSer 297  
Db 567 CAACAACAACATAATAATAATAATGATAATAATAATAAT 605

RESULT 9  
BE908188/c  
LOCUS  
DEFINITION BE908188 754 bp mRNA linear EST 20-OCT-2000  
601500441F1 NIH\_MGC\_70 Homo sapiens cDNA clone IMAGE:3902123 5',  
mRNA sequence.  
ACCESSION BE908188  
VERSION BE908188.1 GI:10402511  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 754)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM9704 row: b column: 12  
High quality sequence stop: 678.  
Location/Qualifiers

FEATURES  
source  
1..754  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3902123"  
/tissue\_type="epithelioid carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_70"  
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: Sali; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.1 kb. Library constructed by Life  
Technologies."  
BASE COUNT 153 a 231 c 239 g 131 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2.31 Length: 754  
Score: 106.50 Matches: 56  
Percent Similarity: 36.65% Conservative: 36  
Best Local Similarity: 22.31% Mismatches: 97  
Query Match: 3.66% Indels: 62  
DB: 10 Gaps: 8

US-09-807-802A-17 (1-534) x BE908188 (1-754)

QY 249 AsnGlnSerGlySerAlaGlnAsnLysAspLeuPheSerArgGlySerProAlaGly 268  
Db 594 CATAGAAGCGAGGGGTTTCCGGGGCCCACTCCCTGCATGGCCGACGCCACCCGGCTG 535  
QY 269 MetSerValGlnProLysAsnTrpLeuProGlyProCysTyrArgGlnArgValSer 288  
Db 534 CTGCCACTACAGCCGGCTCACTGGGCTCCAGGGGACCACGCTCCAGCCATACAGTTTCAT 475  
QY 289 LysThrLysThrAspAsnAsnAsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsn 308  
Db 474 CCTGTGCCAGGTCCCTCTGGTCAAAAGCTCTTGTGTGATGGCCAGG----- 430  
QY 309 LeuAsnGlyArgGluSerIleAsnProGlyThrAlaMetAlaSerHis-LysAspAs 328  
Db 429 -----CCTGGCACAGCTTGAGGGGCCACTGCAGAG-- 399  
QY 328 pGluAspLysPhePheProMetSerGlyValMetIlePheGlyLysGluSerAlaGlyAl 348  
Db 398 -----CTCGCGGTCCCTTGAGGACACGGCTGGCGG 370  
QY 348 aSerAsnThrAlaLeuAspAsnValMetIleThrAspGluGluGluIleLysAlaThrAs 368  
Db 369 TGCC-----TTTCCGGAGGGTGTGCCAAGCTGACGCA 337  
QY 368 nProValAlaThrGluArgPheGlyThrValAlaValAsnPheGlnSerSerThrAs 388  
Db 336 TGGTCTCCAGACC-----TTCCAGGGGCTCCAGTGTCCGACGAAGACGCAAGCCCCGT 283  
QY 388 pProAlaThrGlyAspValHisAlaMetGlyAlaLeuProGlyMetValTrpGlnAspAr 408  
Db 282 CCCCTTGCTGGAGATGTTTCATGCA-----GAAGTACATGCCGTAAA 241  
QY 408 gAspValTyrLeuGlnGlyProIleTrpAlaLysIleProHisThrAspGlyHisPheH 428  
Db 240 TGATGTACACGTACAGGTCCCGGCTTCATGAACATGCCTCGGTTGCGGGGGTCTGCC 181  
QY 428 sProSerProLeuMetGlyGlyPheGlyLeuLysAsnProProGlnIleLeuIleLy 448  
Db 180 GGCCACCCCTTGAGTGGCGGCTTCATCCTCTGG-----CCCCAGGTATGCCCTC-- 132  
QY 448 sAsnThrProValProAlaAsnProProAlaGluPheSerAlaThrLysPheAlaSerPh 468  
Db 131 -----GGTCTCCAGATCGGCGCTCGGAGTTCGTGTGCCAT 97  
QY 468 e-----IleThrGlnTyrSerThrGlyGlnValSerValGluIl 481  
Db 96 TAGGAAGTCGCCGACTAGGACCTGTCCCAAGAAATGCCGGGCCAGGGGACTGCCGGCT 37  
QY 481 eGluTrpGluLeuGlnLysGluAsnSerLys 491  
Db 36 GGTCGAAGAACTCCAAACCCAGTCGGGTAAG 6

RESULT 10  
BE733520/c  
LOCUS  
DEFINITION BE733520 786 bp mRNA linear EST 15-SEP-2000  
601566055F1 NIH\_MGC\_21 Homo sapiens cDNA clone IMAGE:3840973 5',  
mRNA sequence.  
ACCESSION BE733520  
VERSION BE733520.1 GI:10147512  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 786)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)



DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
Plate: LCM531 row: n column: 14  
High quality sequence start: 11  
High quality sequence stop: 772.  
Location/Qualifiers

FEATURES

source  
1..786  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3840973"  
/tissue\_type="choriocarcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_21"

/note="Organ: placenta; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 160 a 236 c 249 g 141 t

ORIGIN

Alignment Scores:  
Pred. No.: 2.46 Length: 786  
Score: 106.50 Matches: 56  
Percent Similarity: 36.65% Conservative: 36  
Best Local Similarity: 22.31% Mismatches: 97  
Query Match: 3.66% Indels: 62  
DB: 10 Gaps: 8

US-09-807-802A-17 (1-534) x BE733520 (1-786)

Qy 249 AsnGlnSerGlySerAlaGlnAsnLysAspLeuLeuPheSerArgGlySerProAlaGly 268  
Db 665 CATAGAGCGGAGGGGTTTCCGGGCCCACTCCCTGCATGGCGACGCCACCCGGGCTG 606  
Qy 269 MetSerValGlnProLysAsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSer 288  
Db 605 CTGCCACTACAGCGGCTCACTGGGCTCCAGGGACCAACGCTCCAGCCATACAGCTTCAT 546  
Qy 289 LysThrLysThrAspAsnAsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsn 308  
Db 545 CCTGTGCCAGGTCCTCTGGTCAAAGCTCTTGTGTATGGCCAGG----- 501  
Qy 309 LeuAsnGlyArgGluSerIleIleAsnProGlyThrAlaMetAlaSerHis-LysAspAs 328  
Db 500 -----CCTGGCACAGCTTGGAGGGGCCACTGCAGAG-- 470  
Qy 328 pGluAspLysPhePheProMetSerGlyValMetIlePheGlyLysGluSerAlaGlyAl 348  
Db 469 -----CTCGCGTCTTCAGGACACGGCTGGCGG 441  
Qy 348 aSerAsnThrAlaLeuAspAsnValMetIleThrAspGluGluIleLysAlaThrAs 368  
Db 440 TGCC-----TTTCCGAGGGTGTGCGAAGCTGACGCA 408  
Qy 368 nProValAlaThrGluArgPheGlyThrValAlaValAsnPheGlnSerSerSerThrAs 388  
Db 407 TGGTCTCCAGACC-----TTCCAGGGGTCCAGTGTCCGACGACGCAAGCCCGT 354  
Qy 388 pProAlaThrGlyAspValHisAlaMetGlyAlaLeuProGlyMetValTrpGlnAspar 408  
Db 353 CCCCTGGCTGGAGATGTTTCATGCA-----GAAGTACATGCCCGTAAA 312  
Qy 408 gAspValTyrLeuGlnGlyProIleTrpAlaLysIleProHisThrAspGlyHisPheHi 428  
Db 311 TGATGTACACGTACAGGGTCCCGGCTTCATGAACATGCCTCGGTTCGGGGGTCTGCC 252  
Qy 428 sProSerProLeuMetGlyGlyPheGlyLeuLysAsnProProGlnIleLeuIleLeu 448

Db 251 GGCCACCCCTTGAGTGGGGCTTCATCTCTGG-----CCCAGGTATGCCTC-- 203  
Qy 448 sAsnThrProValProAlaAsnProAlaGluPheSerAlaThrLysPheAlaSerPh 468  
Db 202 -----GGTCTCCACGATCGGGCCTCGGAGTTCTGTGCCAT 168  
Qy 468 e-----IleThrGlnTyrSerThrGlyGlnValSerValGluIl 481  
Db 167 TAGGAAGTCGCCGACTAGGACCTGTCCAGAAATGCCCGGCCAGGGGACTGCCGGCT 108  
Qy 481 eGluTrpGluLeuGlnLysGluAsnSerLys 491  
Db 107 GGTCTGAAGAACTCCCAACCCCACTGTCGGGTAAG 77

RESULT 11

BQ888825/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 830)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-f@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LCM2314 row: k column: 07

High quality sequence stop: 642.

Location/Qualifiers

1..830

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6083838"

/tissue\_type="ductal carcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_110"

/note="Organ: pancreas; Vector: pOTB7; Site\_1: XhoI;

Site\_2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Library constructed by

Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH\_MGC Library."

BASE COUNT 168 a 248 c 248 g 145 t 21 others

ORIGIN

Alignment Scores:

Pred. No.: 2.67 Length: 830

Score: 106.50 Matches: 56

Percent Similarity: 36.65% Conservative: 36

Best Local Similarity: 22.31% Mismatches: 97

Query Match: 3.66% Indels: 62

DB: 13 Gaps: 8

US-09-807-802A-17 (1-534) x BQ888825 (1-830)

Qy 249 AsnGlnSerGlySerAlaGlnAsnLysAspLeuLeuPheSerArgGlySerProAlaGly 268

```
Db      676 CATAGAAGCGGAGGGTTTCCGGGCCCACTCCCTCGATGGCCGACGCCACCCGGGCTG 617
Qy      269 MetSerValGlnProLysAsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSer 288
Db      616 CTGCCACTACAGCCGGCTCACTGGCTCCAGGGACCAAGCTCCAGCCATACAGTTTCAT 557
Qy      289 LysThrLysThrAspAsnAsnAsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsn 308
Db      556 CCTGTGCCAGGTCCTCTGTGTCAAAGCTCTTGTGTGATGSCCAGGG----- 512
Qy      309 LeuAsnGlyArgGluSerIleIleAsnProGlyThrAlaMetAlaSerHis-LysAspAs 328
Db      511 -----CCTGGCACAGCTGTGGAGGGGCCACTGCAGAG-- 481
Qy      328 pGluAspLysPhePheProMetSerGlyValMetIlePheGlyLysGluSerAlaGlyAl 348
Db      480 -----CTCGCGGTCTTGTGAGGACACGCGGCTGGCGG 452
Qy      348 aSerAsnThrAlaLeuAspAsnValMetIleThrAspGluGluGluIleLysAlaThrAs 368
Db      451 TGCC-----TTTCCGGAGGGTGTGCGAAGCTGACGCA 419
Qy      368 nProValAlaThrGluArgPheGlyThrValAlaValAsnPheGlnSerSerSerThrAs 388
Db      418 TGGTCTCCAGACC-----TTCCAGGGGCTCCAGTGTTCGACGACAGCGCAAGCCCCGT 365
Qy      388 pProAlaThrGlyAspValHisAlaMetGlyAlaLeuProGlyMetValTrpGlnAspAr 408
Db      364 CCCCCTGGCTGGAGATGTTTCATGCA-----GAAGTACATGCGGTAA 323
Qy      408 gAspValTyrLeuGlnGlyProIleTrpAlaLysIleProHisThrAspGlyHisPheHi 428
Db      322 TGATGTACACGTACAGGGTCCCGGGCTTCATGAACATGCTCGGTGCGGGGGTCTGCC 263
Qy      428 sProSerProLeuMetGlyGlyPheGlyLeuLysAsnProProProGlnIleLeuIleLy 448
Db      262 GGCCACCCCTTGAGTGGGGGCTTCATCCTCTGG-----CCCCAGGTATGCTC-- 214
Qy      448 sAsnThrProValProAlaAsnProProAlaGluPheSerAlaThrLysPheAlaSerPh 468
Db      213 -----GGTCTCCACGATGCGGCCTCGGAGTCTCGGCAT 179
Qy      468 e-----IleThrGlnTyrSerThrGlyGlnValSerValGluI 481
Db      178 TAGGAAGTCGCCGGACTAGGACCTGTCCAGAAATGCCCGGCCAGGGGGACTGCCGGCT 119
Qy      481 eGluTrpGluLeuGlnLysGluAsnSerLys 491
Db      118 GGTCGAAGAACTCCAACCCCGAGTCGGGTAA 88
```

```
RESULT 12
BQ682489/c
LOCUS
DEFINITION BQ682489 8034509 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6213998
5', mRNA sequence.
ACCESSION BQ682489
VERSION BQ682489.1 GI:21795168
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 986)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LLCM2380 row: j column: 15  
High quality sequence stop: 612.  
Location/Qualifiers  
1. .986  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6213998"  
/tissue\_type="melanotic melanoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_112"  
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."

BASE COUNT 221 a 297 c 298 g 170 t  
ORIGIN

Alignment Scores: 3.47 Length: 986  
Pred. No.: 106.50 Matches: 56  
Score: 36.65% Conservative: 36  
Percent Similarity: 22.31% Mismatches: 97  
Best Local Similarity: 3.66% Indels: 62  
Query Match: 13 Gaps: 8  
DB:

US-09-807-802A-17 (1-534) x BQ682489 (1-986)

```
Qy      249 AsnGlnSerGlySerAlaGlnAsnLysAspLeuLeuPheSerArgGlySerProAlaGly 268
Db      733 CATAGAAGCGGAGGGTTTCCGGGCCCACTCCCTCGATGGCCGACGCCACCCGGGCTG 674
Qy      269 MetSerValGlnProLysAsnTrpLeuProGlyProCysTyrArgGlnGlnArgValSer 288
Db      673 CTGCCACTACAGCCGGCTCACTGGGCTCCAGGGACCAAGCTCCAGCCATACAGCTTCAT 614
Qy      289 LysThrLysThrAspAsnAsnAsnSerAsnPheThrTrpThrGlyAlaSerLysTyrAsn 308
Db      613 CCTGTGCCAGGTCCTCTGTGTCAAAGCTCTTGTGTGATGGCCAGG----- 569
Qy      309 LeuAsnGlyArgGluSerIleIleAsnProGlyThrAlaMetAlaSerHis-LysAspAs 328
Db      568 -----CCTGGCACAGCTTGGAGGGGCCACTGCAGAG-- 538
Qy      328 pGluAspLysPhePheProMetSerGlyValMetIlePheGlyLysGluSerAlaGlyAl 348
Db      537 -----CTCGCGGTCTTGTGAGGACACGCGTGGCGG 509
Qy      348 aSerAsnThrAlaLeuAspAsnValMetIleThrAspGluGluGluIleLysAlaThrAs 368
Db      508 TGCC-----TTTCCGGAGGGTGTGCGAAGCTGACGCA 476
Qy      368 nProValAlaThrGluArgPheGlyThrValAlaValAsnPheGlnSerSerSerThrAs 388
Db      475 TGGTCTCCAGACC-----TTCCAGGGGCTCCAGTGTTCGACGCAAGACGCAAGCCCCGT 422
Qy      388 pProAlaThrGlyAspValHisAlaMetGlyAlaLeuProGlyMetValTrpGlnAspAr 408
Db      421 CCCCCTGGCTGGAGATGTTTCATGCA-----GAAGTACATGCGGTAA 380
Qy      408 gAspValTyrLeuGlnGlyProIleTrpAlaLysIleProHisThrAspGlyHisPheHi 428
Db      379 TGATGTACACGTACAGGGTCCCGGGTTCATGAACATGCTCGGTGCGGGGGTCTGCC 320
Qy      428 sProSerProLeuMetGlyGlyPheGlyLeuLysAsnProProProGlnIleLeuIleLy 448
```

```
Db 319 GGCCACCCCTTGAGTGGCGGCTTCATCCTCTGG-----CCCCAGGTATGCCTC-- 271
Qy 448 sAsnThrProValProAlaAsnProProAlaGluPheSerAlaThrLysPheAlaSerPh 468
Db 270 -----GGTCTCCACGATCGGCGCTCGGAGTTCTGTGCCAT 236
Qy 468 e-----IleThrGlnTyrSerThrGlyGlnValSerValGluIl 481
Db 235 TAGGAAGTCGCGGACTAGGACCTGTCCAGAAATGCCGGGCCAGGGGACTGCCGGCT 176
Qy 481 eGluTrpGluLeuGlnLysGluAsnSerLys 491
Db 175 GGTGGAAGAACTCCAACCCAGTCGGGTAAG 145

RESULT 13
BU800802
LOCUS SJF2CVC08 SJF Schistosoma japonicum cDNA, mRNA linear EST 12-FEB-2003
DEFINITION BU800802
ACCESSION BU800802
VERSION BU800802.1 GI:28357809
KEYWORDS EST.
SOURCE Schistosoma japonicum
ORGANISM Schistosoma japonicum
REFERENCE Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
AUTHORS Strigeidida; Schistosomatidae; Schistosomatidae; Schistosoma.
1 (bases 1 to 606)
Hu,W., Yan,Q., Shen,D., Liu,F., Xu,X., Zhu,Z., Zhang,X., Wang,J.,
Sun,J., Xu,X., Wang,Z., Zeng,L., Rong,Y., Wu,X., Qu,J., Xu,Z.,
Huang,J., Ma,Y., Wang,S., Wang,Z., Xue,C., Feng,Z., Chen,Z. and Han
,Z.
Expressed sequence tags from female adults of Schistosoma japonicum
Unpublished
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn.
Location/Qualifiers
1..606
/organism="Schistosoma japonicum"
/mol_type="mRNA"
/db_xref="taxon:6182"
/sex="female"
/tissue type="Whole body"
/dev_stage="adult"
/lab_host="rabbits"
/clone_lib="SJF"
BASE COUNT 242 a 85 c 40 g 230 t 5 others
ORIGIN

Alignment Scores:
Pred. No.: 2.37 Length: 606
Score: 105.00 Matches: 51
Percent Similarity: 36.29% Conservative: 35
Best Local Similarity: 21.52% Mismatches: 61
Query Match: 3.61% Indels: 90
DB: 13 Gaps: 14

US-09-807-802A-17 (1-534) x BU800802 (1-606)

Qy 111 AsnPheLysLeuPheAsnIleGlnValLysGluValThrThrAsnAspGlyValThrThr 130
Db 18 AACTATAATATCTCAATCTTCTTTA---CATACATCATCGTCATCATCAATAACA 74
Qy 131 IleAlaAsnAsnLeuThrSerThrValGlnValPheSerAspSerGluTyrGlnLeu--- 149
Db 75 ATAGTACCATCATTCATCATCT-----TATTCATTCACCGTTTAGTTAAAT 122
Qy 150 -----ProTyrValLeuGlySerAlaHisGlnGlyCysLeuProProPheProAlaAsp 167
Db 123 CCATTTCATATATATAAATTTTCAAAAAAATCT-----TCAGAT 164
```

```
Qy 168 ValPheMetIleProGlnTyrGlyTyrLeuThrLeuAsnAsnGlySer----- 183
Db 165 TTTTATTATTAAACAAAAAATCAAATTATTCTACAAATGGTTTCATATCATNNNNA 224
Qy 184 -----GlnAlaValGlyArgSerSerPheTyrCysLeuGlu----- 195
Db 225 GAATTAATACGTACATTACAAAGTCGTAAATCATATTGGTGTTCACAAAATTTGAATT 284
Qy 196 -----TyrPhePro-----SerGln 200
Db 285 TTTGCACCATTTTTCATTTTATTATTTTCCATTATTTATGATTATTTCTACACAT 344
Qy 201 MetLeuArgThrGlyAsnAsnPheThrPheSerTyrThrPheGlu----- 215
Db 345 TTATTATCACTAATTAATAAAATTGAAATTTTCTACAAACATTTAAATGATTTTATGTGAT 404
Qy 216 -----GluValProPheHisSerSerTyrAlaHis 225
Db 405 TTACCAGAAATTGAAATATTGAATCGTGCACGTATACGATATCATCGT----- 452
Qy 226 SerGlnSerLeuAspArgLeuMetAsnProLeuIleAspGlnTyrLeu---TyrTyrLeu 244
Db 453 -----CAACAATATGTAAACATATGTT 476
Qy 245 AsnArgThrGlnAsnGlnSerGlySerAlaGlnAsn-----LysAspLeuLeuPhe 261
Db 477 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 536
Qy 262 SerArgGlySerProAlaGlyMetSerValGlnProLysAsnTrpLeuProGlyProCys 281
Db 537 TCACAT---CCTCAAGCAACACATTTCTGTAAT----- 566
Qy 282 TyrArgGlnGlnArgValSerLysThrLysThrAspAsnAsnSerAsn 298
Db 567 TATCAACACCAACAA-----CAACATAATAATAATAATAATAATAATAATAATAATA 605

RESULT 14
BI253826/c
LOCUS BI253826 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5115479 5',
DEFINITION BI253826 775 bp mRNA linear EST 17-JUL-2001
ACCESSION BI253826 mRNA sequence.
VERSION BI253826.1 GI:14805634
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11280 row: n column: 24
High quality sequence stop: 701.
Location/Qualifiers
1..775
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5115479"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_lib="NIH_MGC_12"
FEATURES
source
```





Qy 416 IleTrpAlaLysIleProHisThrAspGlyHisPheHisProSerProLeuMetGlyGly 435  
: : : : :  
Db 337 ATGTGGATGTTGCTTTTGACCTACGGGGGGCATCTGAGTCCAGCTCCCCCAAGATGAGC 396  
Qy 436 PheGly-----LeuLysAsnProProProGlnIleLeuIle 447  
: : : : :  
Db 397 TGCAGCCCCCAGAGAGAGCTCTGCACGTCACCAAGTAACAGGCCCCAGCCTCCAGGCC 456  
Qy 448 LysAsnThr-----ProValProAlaAsnProProAlaGluPheSerAlaThrLysPhe 465  
: : : : :  
Db 457 CCCAACTCCGCCCGCAGCCTCTCCCGCTCTGGATCCTGCACCTTAACACTCGACTCTGCTG 516  
Qy 466 AlaSerPheIleThrGln-----TyrSerThrGiyGlnValSerValGluIleGlu 482  
: : : : :  
Db 517 CTCATGGGAAGAACAGAATTGCTCCTGTCATGCAACTAAT-----TCAATAAAACTGTCT 570  
Qy 483 TrpGluLeuGlnLysGluAsnSerLysArgTrpAsnProGlu 496  
: : : : :  
Db 571 TGTGAGCTGAAAAAAGAAAAAAGAAAAAAGAGAGAACTCGGAG 612

Search completed: February 15, 2004, 01:29:26  
Job time : 2990.43 secs